Applicants: Graham P. Allaway et al.
Serial No.: 09/904,356
Filed: July 12, 2001
Exhibit 16

Fundamental Immunology, Fourth Edition,
edited by William E. Paul
Lippincott-Raven Publishers, Philadelphia © 1999.

CHAPTER 3

Immunoglobulins: Structure and Function

J. Kimble Frazer and J. Donald Capra

INTRODUCTION

Immunoglobulin is the crux of the humoral immune response. As a cell surface receptor on B lymphocytes, immunoglobulin is responsible for instigating cellular processes as diverse as activation, differentiation, and even programmed cell death. As secreted antibody in plasma and other bodily fluids, immunoglobulin is able to bind foreign antigen, thereby either neutralizing it directly or initiating steps necessary to arm and recruit effector systems such as complement or antibody-dependent cell cytolysis by monocytic phagocytes. The ability of immunoglobulin to perform such a wide array of duties can be attributed to evolution's clever usage of a structural paradigm—the immunoglobulin domain—and its duplication, diversification, and elaboration upon that design to endow it with an assortment of functional qualities.

Despite the variety of purposes served by immunoglobulin molecules, one feature remains common to virtually all considerations of immunoglobulin structure and function: immunoglobulins have an amazing capacity to interact with other molecules. In one sense, immunoglobulins must be able to effectively bind a finite set of invariant partners, such as Fc receptors, signal-transducing molecules, and components of the complement cascade. In another sense, immunoglobulins, collectively, must meet the challenge of being able to recognize an essentially infinite array of antigenic determinants. More remarkable, perhaps, is the fact that immunoglobulin is frequently called upon to fulfill both of these binding responsibilities simultaneously, and in such a way as to mediate significant biological effects. As such, immunoglobulin molecules may be viewed as a marriage between the constraints engendered by biological continuity and the quest for diversity superimposed upon this evolutionary framework.

The lengths to which evolution has gone in order to bestow immunoglobulin with these conflicting capabilities has been the subject of intense scientific scrutiny, and has yielded innumerable fascinating insights into immunology, genetics, protein chemistry, and the discipline of biology as a whole. In trying to understand how antibody is able to recognize such a multitude of different specificities, science has benefited from the discovery of both VDJ recombination (see Chapter 5) and somatic hypermutation (see Chapter 25). In an attempt to reconcile the incongruity entailed by the observation of highly divergent N-terminal regions coupled to constant C-terminal domains, research has gained not only the once-heretical "two genes, one polypeptide" hypothesis (1), but also the concept of isotype switching (see Chapter 24). Thus, studies into immunoglobulin diversity have proven to be extremely profitable scientific endeavors. In addition, while diversity has been a hallmark of the study of immunoglobulin since it was first recognized to be a salient feature, several aspects that derive from immunoglobulins' underlying uniformity have been used to glean understanding into protein structure–function relationships in general.

Immunoglobulins were the first molecules described from the ancestral immunoglobulin superfamily (IgSF) (2–4). As an ever-

J. K. Frazer: Graduate Program in Molecular Microbiology and Medical Scientist Training Program, University of Texas Southwestern Medical School, Dallas, TX 75235-9140.
J. D. Capra: President and Scientific Director, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104.

expanding gene family, members of the IgSF have been shown to be vital to issues of cell–cell interaction and molecular recognition in a variety of cell types and across several taxonomic boundaries. Many molecules central to the functioning of the immune system, including the antigen-specific chains of the T-cell receptor (TCR) (see Chapter 10) and the class I and II major histocompatibility complex (MHC) antigens (see Chapter 8), are counted among this group. Common to all members of the superfamily is the presence of one or more immunoglobulin-like domains. Three-dimensional structural analyses of proteins containing these regions have demonstrated that the conserved amino acid sequences that make up an immunoglobulin homology domain comprise a recurring structural motif that can fold into a compact globular subunit. These subunits, in turn, are capable of integrating into complex macromolecules (5,6). As a result, different immunoglobulin molecular structures are similar not only to each other, but also to a multitude of other important proteins.

Because such a concerted scientific effort has been made to understand the way in which immunoglobulin functions, a large volume of sequence information—at both the nucleotide and amino acid levels—is available in both the literature and public databases. Indeed, immunoglobulins have likely been sequenced more frequently than any other class of gene or protein. Similarly, immunoglobulins have been well represented in structural studies, crystallographic and otherwise, to an unprecedented degree. This mass of work has allowed a number of conclusions regarding structure–function relationships of immunoglobulins to be made. Specifically, the aim of this chapter is not to compile an exhaustive catalogue of all extant work on the topic of immunoglobulin sequences, but rather to present the essential features of immunoglobulin structure and their relation to immunologic function as is currently understood. Further, because immunoglobulin proteins have been so evolutionarily valuable, they can be found, in one form or another, throughout vertebrate species. Many of these molecules are only now being characterized, and surely many more are yet to be identified. As a consequence of this diversity, however, it is impossible to relate all of the details of immunoglobulin structure and function in their entirety. Instead, unless otherwise noted, the examples of human and murine immunoglobulins will be used as models to convey the general conclusions garnered from scientific insight and experimentation into this important and fascinating class of proteins. The organization of this discussion will begin, following an introduction to basic immunoglobulin features, with a consideration of the primary structure of antibody molecules and proceed through the secondary, tertiary, quaternary, and higher order immunoglobulin structural topics that derive from its sequence. Once this foundation has been laid, the functional attributes of immunoglobulin will be considered, with an eye to correlating an antibody's capacities—to the extent which it is possible—with that of its structure. A section on the IgSF follows, which will briefly address its evolution and also specifically detail particular IgSF members critical to immune responses that are not explicitly covered elsewhere in this volume.

GENERAL IMMUNOGLOBULIN STRUCTURE, NOMENCLATURE, AND HISTORY

Structural Considerations

Figure 1 presents a diagrammatic representation of an antibody molecule. The typical immunoglobulin monomer is comprised of four polypeptide chains complexed together via hydrophobic interactions and stabilized by disulfide bonds. Due to allelic exclusion (see Chapters 5 and 6) B lymphocytes usually express only one functionally rearranged heavy chain gene and only one light chain polypeptide as well. Consequently, complete immunoglobulin proteins are composed of two identical heavy chain polypeptides of approximately 55 kD and two identical light chains of 25 kD. Each heavy and light chain pair is joined by one or more interchain disulfide bonds, and also relies upon noncovalent interactions to properly orient the two chains relative to each other. One such "half-antibody" contains a single antigen binding site (i.e., it is monovalent). The complete four polypeptide chain monomer is formed by similar hydrophobic bonding between the two heavy chains, and it also utilizes one or more disulfide bonds to stabilize the complex. Thus, a complete immunoglobulin molecule is bivalent with two identical sites for potential binding of antigen. As such, an immunoglobulin may be thought of as a "dimer of a heterodimer," although these half-molecules do not occur naturally.

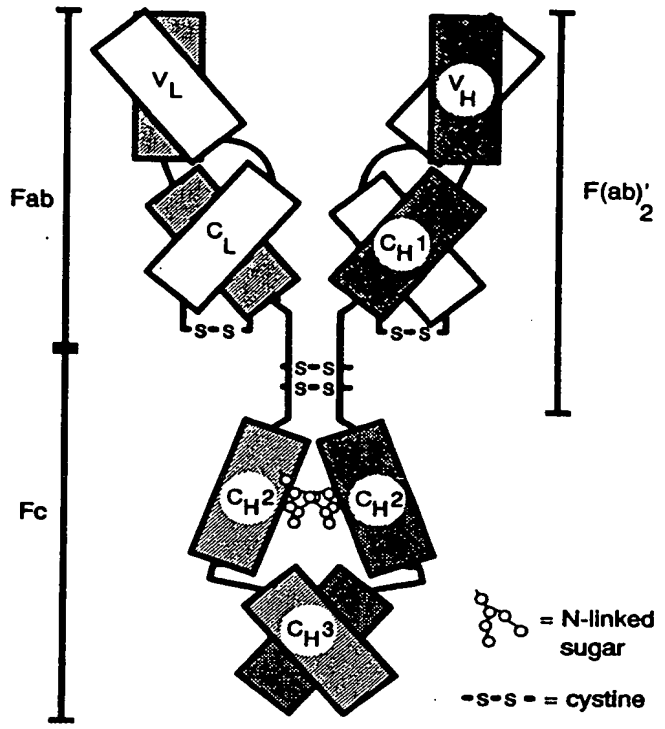


FIG. 1. Schematic representation of a prototypic immunoglobulin monomer. Each box symbolizes a complete immunoglobulin domain from either the heavy (*shaded boxes*) or light (*unshaded*) chain. Labeling of domains follows standard nomenclature, as outlined in the text. Interchain disulfide bonds are denoted by *black bars*. Note that these bonds are present between both heavy and light chain pairs and between the two heavy chains. Conserved N-linked carbohydrate occurs on all C_H2 domains as shown, although some immunoglobulins are also glycosylated at additional sites elsewhere in the molecule. Also of note is the fact that all of the domains associate to form dimeric modules (V_H/V_L, C_H1/C_L, and C_H3/C_H3), except C_H2 domains. The Fab, Fc, and $F(ab)'_2$ proteolytic fragments are demarcated by *bars* to either side of the diagram. (From ref. 6a, with permission.)

Each individual polypeptide chain consists of two to five domains of approximately 110 amino acids (7), each capable of folding independently. These domains form compact, protease-resistant structures which serve as the fundamental unit of immunoglobulin structure. The interactions that allow for the formation of the aforementioned immunoglobulin monomer almost exclusively occur in pair-wise fashion between domains of two different polypeptide chains (see Fig. 1), such that the functional modules of an antibody are in fact dimerized domains. In addition, as each domain of an antibody molecule is encoded by a separate exon, immunoglobulin domains also serve as the essential element of antibody genetics. In this light, it is easy to recognize how evolution has used the prototypical immunoglobulin domain as a substrate for experimentation, and as a result different domains have attained distinct structural and functional attributes. Moreover, the presence of one or more "immunoglobulin homology domains" also proves to be the distinguishing characteristic for inclusion in the immunoglobulin gene superfamily. Thus, the duplication and adaptation of the Ig homology domain has occurred not only within the context of formal "immunoglobulin genes," but also in the greater scope of the IgSF, which far predates the emergence of antibody. In either case, the archetypal immunoglobulin domain has clearly proven to be a powerful evolutionary tool, as will be detailed below and in greater detail throughout this chapter.

The hallmark of all Ig domains is the presence of a structural motif termed the *immunoglobulin fold*. This characteristic feature is actually a specialized "β-barrel" typically comprised of seven polypeptide strands, which form antiparallel β-pleated sheets in the folded domain. This configuration is depicted in Fig. 2, which was deduced from x-ray diffraction studies of an immunoglobulin light chain (8). Each Ig domain is composed of two β-pleated sheets, one containing four β strands, the other consisting of at least three β strands (represented by arrows in Fig. 2). Loops of variable length connect the different strands, allowing the β sheets to form. The two β-pleated layers are oriented in a sandwich, enclosing a hydrophobic interior. Further stability is provided by a disulfide bond near the domain's core, which covalently links the two sheet layers. The cysteines that contribute this bond are conserved in all immunoglobulins, and in almost all proteins that possess Ig-like domains. Two residues, a tryptophan in strand 3-1 and an aromatic residue that precedes the second half-cystine, are also maintained consistently and serve to protect the disulfide bond in the three-dimensional structure. Other conserved features include hydrophobic core residues, which stabilize the inside of the sandwich, and glycine and proline loop residues, which provide the flexibility necessary for the formation of these interconnecting sequences (9–12).

Since the hydrophobic core residues are predominantly responsible for promoting the folding of the β sheets, and thus the entire immunoglobulin fold, the sequences of the loop residues are free to vary considerably. This, in turn, grants loop residues the freedom to serve as substrates for selection, at the level of selection of a particular antibody in an immune response and at the level of natural selection in phylogeny. In this way, the prototypical immunoglobulin homology domain serves as a potent cofactor for the evolution of both organismal immunity and that of the species in general.

Immunoglobulin Nomenclature

Light chains contain two such immunoglobulin domains, whereas a heavy chain is made up of either four or five domains, depending on the type of heavy chain (isotype) used by the antibody in question. Different immunoglobulin domains possess different structural and functional characteristics, and their naming, in part, reflects these differences. The amino-terminal domain of each chain, whether of the heavy or light type, is termed a *variable* (V) region due to the discovery of extensive sequence divergence between different antibody proteins in this part of the molecule. These are designated V_H and V_L for heavy and light chains, respectively. V regions have been demonstrated to be responsible for the antigenic specificity of the immunoglobulin.

Carboxy-terminal domains, on the other hand, display considerably less sequence variation within a given isotype and are referred to as *constant* (C) regions. Heavy chain C regions are numbered C_H1, C_H2, and so on, beginning with the most V region-proximal domain. The constant region domains of the heavy chain have been shown to be responsible for many aspects of antibody function, including interaction with Fc receptors, complement fixation, transplacental transfer, the ability to multimerize, and the capacity to be secreted on mucosal surfaces. Because different heavy chain isotypes have different C region domains (i.e., the C_H3 domains of different isotypes are distinct), these capabilities vary with the class of the particular antibody. Five major classes of heavy chain C regions exist: alpha (α), gamma (γ), delta (δ), epsilon (ε), and mu (μ). As a direct consequence of the correlation between the

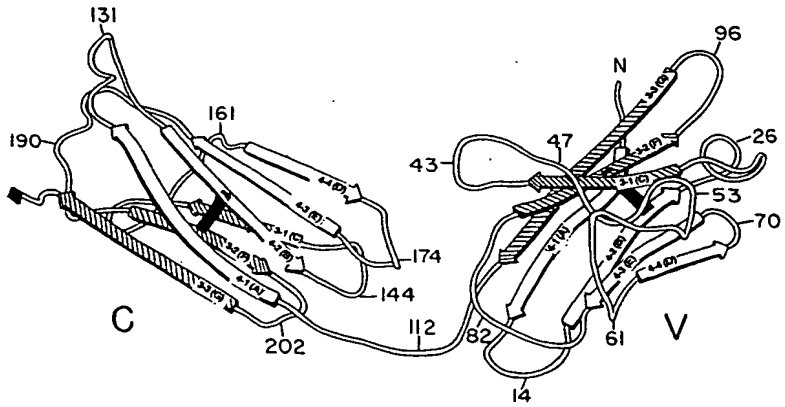


FIG. 2. Ribbon drawing of the V and C domains of a light chain. β strands are depicted as *arrows*, with those of the four-stranded face *unshaded* and those of the three-stranded face *shaded*. Strands are numbered according to Edmundson and lettered (*in parentheses*) according to Hood. Intrachain disulfide bonds are represented as *black bars*. Selected amino acids are numbered, with position 1 being the N-terminus. Residues 26, 53, and 96 correspond to amino acids in CDRs 1, 2, and 3, respectively. The dimerization surfaces of each domain (four-strand side of the C domain, three-strand side of the V domain) face upwards. (Adapted from ref 8, with permission.)

heavy chain class of an antibody and its resultant effector functions, immunoglobulins are named according to their heavy chain, using an English-letter terminology (IgA, IgG, IgD, IgE, IgM), which corresponds to their Greek letter isotypes. Specific domains of C regions are often designated according to the class of heavy chain from whence they originate as well (i.e., the C_H3 domain of a μ antibody is signified by $C_\mu3$). Owing to the propensity of immunoglobulin domains to evolve independent of one another, oftentimes a particular domain of a specific isotype may be responsible for one or more functional characteristics of the entire antibody, making this naming system particularly relevant. On the other hand, the constant regions of light chains, possessing only one C domain, are usually simply denoted by C_L. The two light chain classes, kappa (κ) and lambda (λ), may be indicated by the use of C_κ or C_λ designations. No distinct functional attributes have as yet been ascribed to either the κ or λ light chain isotypes.

Finally, immunoglobulins also have hinge regions located C-terminal to the C_H1 domains of their heavy chains. In the case of heavy chains of the μ and ϵ isotypes, the hinge is so elongated that it is actually an extra immunoglobulin domain, explaining the presence of a fifth C domain in these molecules. Other heavy chain classes use shorter stretches of protein, which are thought nonetheless to have evolved from the $C_{\mu/\epsilon}2$ domain. Consistent with the independent evolution of the other domains of immunoglobulin genes, hinge regions are generally encoded by individual exons as well. As the name implies, the hinge permits a generous degree of flexibility between the antigen-binding and effector-interacting components of the molecule. Thus, the hinge domain facilitates linking of the two disparate elements of immunoglobulin function: the ability to interact with an endless array of structural determinants on antigenic surfaces (mediated by V regions) and the capacity to interact with a limited number of effector-activating molecules (mediated by C regions). In addition, disulfide bridges between the two heavy chains typically occur within the boundaries of the hinge region, allowing the complete tetrameric complex to form.

Hence, in many cases, the discrete elements of immunoglobulin structure defined both genetically and structurally as immunoglobulin domains are also responsible for specific functional qualities. Moreover, in addition to the one-domain-per-exon correlation that exists for immunoglobulins, in both heavy and light chains the V region domain and the domains of the different C regions are in fact distinct genes (13). This type of genetic arrangement allows the ability to recognize a specific antigen to be united with the effector functions that are most appropriate for that particular immune response at that particular time. In this regard, then, it is clear that antibodies truly embody the linkage of structure to that of function.

An Historical Perspective

Long before x-ray diffraction of crystals had yielded the keys to dissecting the structure of immunoglobulin, other seminal studies had been performed that, in retrospect, agree completely with the conclusions drawn from crystallization analysis. Many of these experiments focused on the basic protein chemistry of pooled IgG, using the techniques of proteolysis, reduction, and denaturation. First, it was revealed that papain digestion of IgG would render two types of protein fragments: Fab, a monovalent antigen binding fragment, and Fc, an easily crystallizable fragment (14). Soon after, it was recognized that pepsin treatment of IgG produced an antigen-binding fragment designated $F(ab)'_2$ which had bivalent activity (15). Furthermore, if this fragment was treated with reducing agents, two univalent Fab' fragments could be obtained. These different fragments are schematically represented in Fig. 1. Reduction and dissociation of IgG also demonstrated that identical heavy and light chains were complexed via disulfide bonds (16). This and other work eventually allowed investigators to decipher a working model for immunoglobulin structure consisting of four polypeptide subunits—two identical heavy chains and two identical light chains—stabilized by multiple interchain disulfide bonds (17,18), which we now know to be correct.

Early studies of a different type also proved successful in revealing information about antibodies; these experiments focused upon utilizing the immune system itself as a means to decode aspects of immunoglobulin structure (reviewed in refs. 19–21). As large glycoproteins, antibodies themselves are potent immunogens capable of eliciting vigorous humoral immune responses. Investigators used immunoglobulin preparations (initially either heterogeneous total serum immunoglobulin or homogeneous myeloma or plasmacytoma proteins, and later monoclonal antibodies from hybridomas) as antigens to generate antibody responses by injecting them into animals of differing species or different animals of the same species. The antibodies produced by these immunization protocols proved useful in resolving several key elements of immunoglobulin structure, and many of the antigenic determinants recognized by these antisera have subsequently been shown to correlate exactly with known structural features of immunoglobulins.

A three-tiered serological classification scheme for immunoglobulin was devised using these antisera (after adsorption) as reagents to categorize antibody molecules into distinct groups. The first tier of organization is that of the *isotype*. Isotypes define C region determinants and, as such, distinguish heavy and light chain constant regions from one another. Initially, five heavy chain classes were recognized and given the Greek letter designations mentioned in the previous section. The presence of these five isotypes in virtually all mammals for which immunoglobulin profiles have been determined indicates that the divergence of C region genes occurred at an early stage of mammalian evolution. Similarly, light chain constant regions were also divided into discrete κ and λ classes. Soon after, refinements made clear that two of the human heavy chain classes, α and γ, in fact contained several related members that could be further divided into subclasses. Human IgA is separated into $\alpha1$ and $\alpha2$ subclasses, and human IgG is separated into four γ subclasses: $\gamma1$, $\gamma2$, $\gamma3$, and $\gamma4$. Murine IgG is also composed of four γ subclasses ($\gamma1$, $\gamma2a$, $\gamma2b$, and $\gamma3$), although their structural and functional characteristics—and their abbreviated designations—do not agree with their human counterparts, thus indicating this diversification occurred after these species' evolutionary divergence. Each different isotype (whether class or subclass) is represented by a separate C region gene in the haploid genome, and all isotypes are present in the sera of all normal individuals of a given species.

Allotypes, on the other hand, refer to determinants found on the antibodies of some, but not all, members of a species. These determinants are encoded by one allele of a particular C region gene (either heavy or light chain) and are inherited in typical Mendelian fashion as autosomal dominant traits. Compilations of human allotypes are summarized in ref. 22 and are covered more extensively in ref. 23. Whereas isotypes and allotypes are localized to the C regions of immunoglobulins, *idiotypes* are antigenic determinants found on the V regions of antibodies, and they frequently correlate

with binding specificity. Generally, idiotypes are present only in an individual member of a given species, and these antigenic epitopes reflect the uniqueness of each individual immunoglobulin molecule. An idiotypic determinant defined by a monoclonal antibody is called an *idiotope*. Idiotypes are not always restricted to the individual, however. Occasionally, when two individuals are challenged with the same antigen, they will produce antibodies that share the same idiotypic determinant(s). In such cases, the idiotype is called a *cross-reactive,* or *public, idiotype* (24). Cross-reactive idiotypes represent the usage of the same V gene segment by different individuals. Thus, idiotypes may be best thought of as being restricted not to the individual organism, but to the individual immunoglobulin molecule.

Obviously, a tremendous amount of effort, using a variety of scientific approaches, has been focused upon attempts to understand immunoglobulin structure. It is remarkable, though, that the vast majority of this early work, whether utilizing protein chemistry to resolve basic structural characteristics or manipulating the humoral immune response to generate reagents to classify immunoglobulin proteins relative to one another, has in fact identified, and in many cases answered, many of the crucial questions of immunoglobulin structure correctly. As shall be discussed in the following sections, many of the crucial structural features of the antibody—from primary sequence to quaternary associations—were first inferred from these initial landmark studies.

IMMUNOGLOBULIN STRUCTURE

Primary Structure—Two Genes, One Polypeptide

The assertion that each immunoglobulin chain derived from two distinct genetic entities was a novel and provocative hypothesis at the time it was proposed (1), and it has proven to be correct. Owing to this genetic independence, the structures of the V and C regions of immunoglobulin will be treated separately here as well, although they obviously share numerous commonalties.

Among the most remarkable discrepancies between V and C regions are the differences in their genetic organization. While the different heavy and light chain C regions are encoded by fewer than 20 genes, V region genes (V_H, V_κ, and V_λ) number in the hundreds. Further distinguishing the V region loci is the fact that genes for complete V_H or V_L domains are not present in the genome originally, but are "recombined" at the genetic level according to the processes of somatic diversification described in Chapter 5. This recombination of V, D, and J elements to form functional heavy chain genes (or V and J in the case of light chains) imposes another degree of diversity upon the V regions. Due to these complexities, sequence variability is in fact the hallmark of variable region domains. In addition to the differences between V and C domains in their genetic design and construction, V region sequences also have uniquely identifiable characteristics. One of the most easily recognizable features is the fact that V regions are approximately 16 residues longer than the prototypic 110 amino acid immunoglobulin domain. These extra residues allow V regions to form a distinctive immunoglobulin fold structure using two additional β strands, which distinguishes V domains from C domains and also has implications for V region function. Also, the processes of V(D)J recombination can further alter the germline–encoded length of V regions, subtly affecting V domain structure and function as well.

V regions form the amino-terminal domains of heavy and light chains, and their primary responsibility is the binding of antigen. The promotion of the capacity to recognize antigenic determinants has been the driving force behind the evolution of V genes, both at the structural level of individual genes and for the evolution of the different V loci. When sequence data first became available from antibody proteins (see Fig. 3), it was apparent that great variation existed between V regions relative to that found between C regions. A means was developed to quantitate this variation whereby variability was defined as the number of different amino acids observed at a given position divided by the frequency of the most common amino acid at that position (25). Using this equation, an invariant residue would have a variability equal to one, whereas the theoretical upper limit for a position occupied by each of the 20 amino acids in a random fashion would be 400. This can be illustrated graphically by plotting the variability scores of a particular protein against its residue number, as is demonstrated in Fig. 4. Variability plots of this type established not only that V regions were characterized by diversity in their sequences, but that this variation was principally clustered in three regions, which were deemed *hypervariable regions* (HVRs). It was hypothesized that these highly variable segments of heavy and light chains would coordinate in such a way as to form the antigen-combining site

```
LIGHT CHAIN VARIABLE REGIONS

     1                                          A 27  F
AG:  DIQMTQSPSSLSASVGDRVTITCQASQ------DINHYL
LEN: **V*****NS*AV*L*E*A**N*KS**SVLYSSNSKN**
TI:  E*VL****GT**L*P*E*A*LS*R***S-----VSNSF*

AG:  NWYQQGPKKAPKILIYDASNLETGVPSRFSGSGFGTDFT
LEN: A****K*GQP**L***W**TR*S***D******S*****
TI:  A****K*GQ**RL***V**SRA**I*D******S*****

                                                109
AG:  FTISGLQPEDIATYYCQQYDTLPRTFGQGTKLEIKRT
LEN: L***S**A**V*V******YST*YS**********T
TI:  L***R*E***F*V******GSS*S*******V*L**T

KAPPA LIGHT CHAIN CONSTANT REGIONS

     108
AG:  TVAAPSVFIFPPSNEQLKSGTASVVCLLNNFYPREA
LEN: **********************************
TI:  **********D***********************

AG:  KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT
LEN: **********************************
TI:  **********************************

                                          214
AG:  LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
LEN: **********************************
TI:  **********************************
```

FIG. 3. Amino acid sequences of human antibody light chains. *Dashes* denote gaps introduced to optimally align sequences; *asterisks* represent identity relative to the top sequence. These are among the first immunoglobulin sequences ever obtained, demonstrating the dramatic differences in variation between V and C regions. The appearance of clusters of conserved residues within the different V domains also illustrates the necessity of a system to accurately quantify variations between several sequences (see Fig. 4). (Developed from the sequence compendium of Kabat et al. in ref. 24a.)

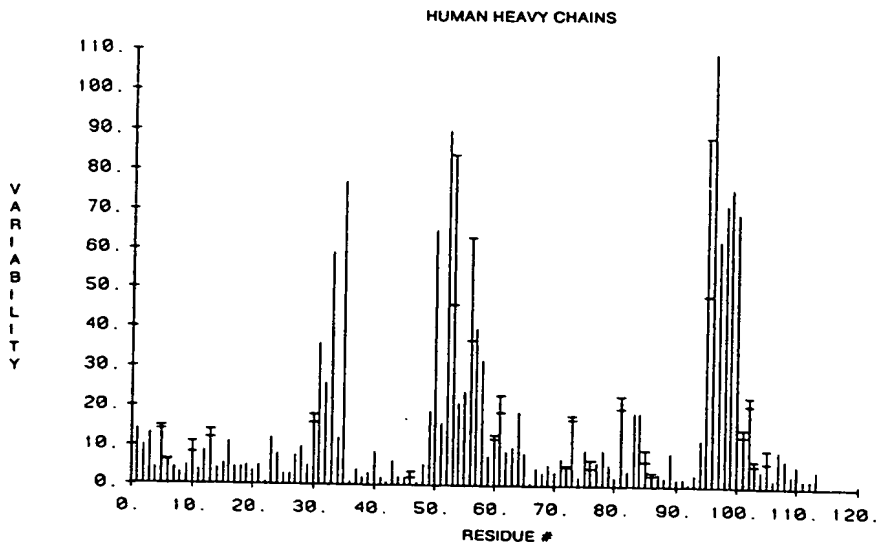


FIG. 4. Variability plot of human heavy chains. The hypervariable regions are apparent as the three obvious peaks in the graph. (From ref. 24a, with permission).

(reviewed in ref. 26); thus they were termed *complementarity-determining regions* (CDRs) as well.

Variability analysis also determined that other stretches of V region sequence were reasonably well conserved from protein to protein; these were presumed to perform basic structural functions necessary for proper folding of all V domains. Accordingly, they were dubbed *framework regions* (FRs) because they provide the platform that supports the CDRs. Structural analyses have confirmed that the FRs largely coincide with the β strands of the immunoglobulin fold, while CDRs, on the other hand, chiefly correspond to the loops that join β strands on the C region-distal end of the V domain. A linear representation of this association is shown in Fig. 5; note that CDRs 1, 2, and 3 join β strands B and C, C and D, and F and G, respectively. Significant sequence motifs are also apparent in this comparison of six V region proteins: a W/F-G-X-G motif in FR4 that is common to all V domains (27), the V_H-specific G-L-E-W-hydrophobic stretch in FR2 (28), and the V_L-specific sequence P-hydrophilic-hydrophobic-L-hydrophobic in the analogous FR2 location (28). These motifs are vital for proper dimerization of domains and will be discussed further in the section on quaternary immunoglobulin structure. Another important distinction between heavy and light chain V region sequences is also apparent in Fig. 5: relative to V_L domains, V_H regions generally utilize longer FR1 and CDR2 segments and shorter CDR1 and FR2 stretches. While the vast number of V genes precludes the ability to definitively assign boundaries for FRs and CDRs that are constant among all immunoglobulin V regions, Table 1 summarizes the traditional positions that delineate these areas for both heavy and light chains.

The presence of hundreds of different germline V region genes obviously contributes greatly to the sequence diversity of different variable domains. However, the somatic process of V(D)J recombination (see Chapter 5) further accentuates V region variability, specifically targeting the CDR3 of the protein (29). In this system, approximately 100 unique V genes (V_H, V_κ, or V_λ loci) encode the N-terminal FR1-CDR1-FR2-CDR2-FR3-5′CDR3 portions of V regions, while four to six "joining" (J) minigenes code for the carboxy-terminal 3′CDR3-FR4 segments. Heavy chains also incorporate one of about 30 short "diversity" (D) gene segments between V and J genes to generate complete V region domains. The relationship between rearranged V(D)J gene segments and the FR/CDR organization of the V region is schematically represented in Fig. 6. The combinatorial assortment of gene segments to form complete heavy and light chains, followed by the combinatorial assortment of heavy and light chains with each other to form antigen-binding V_H/V_L dimers, results in a practically limitless number of V domain structures. Moreover, during the recombination process itself, the activity of exonucleases and untemplated N-segment additions (29), templated P-nucleotide incorporation (30), and D-D fusion events (31) can boost the diversity of CDR3 even further. Finally, superimposed on these aspects of "combinatorial" and "junctional" diversity-generating mechanisms, somatic hypermutation (32,33; see Chapter 25) serves to introduce still more variation by altering residues throughout the V region.

Despite—and perhaps as a result of—the seemingly endless number of possible V region sequences, sophisticated schema have emerged for their classification. These groupings are based upon homology-based hierarchies that directly reflect the evolution of the antibody gene loci. Members of a group are more similar to each other than to all other sequences from other groups and share linked amino acid substitution patterns, which serve as "identifiers" for the various classifications. The most evolutionary distant grouping is, of course, that of the V regions themselves, followed by the V_H, V_κ, and V_λ distinctions, which represent the separate V gene loci. In humans, the heavy chain locus is found on chromosome 14, and the κ and λ loci are found on chromosomes 2 and 22, respectively (34–36). In the mouse, these genes are located on chromosomes 12, 6, and 16 (37–39). Other stratifications for V region organization also mirror the evolution of the antibody gene loci. The use of "clans" to categorize V genes has demonstrated the

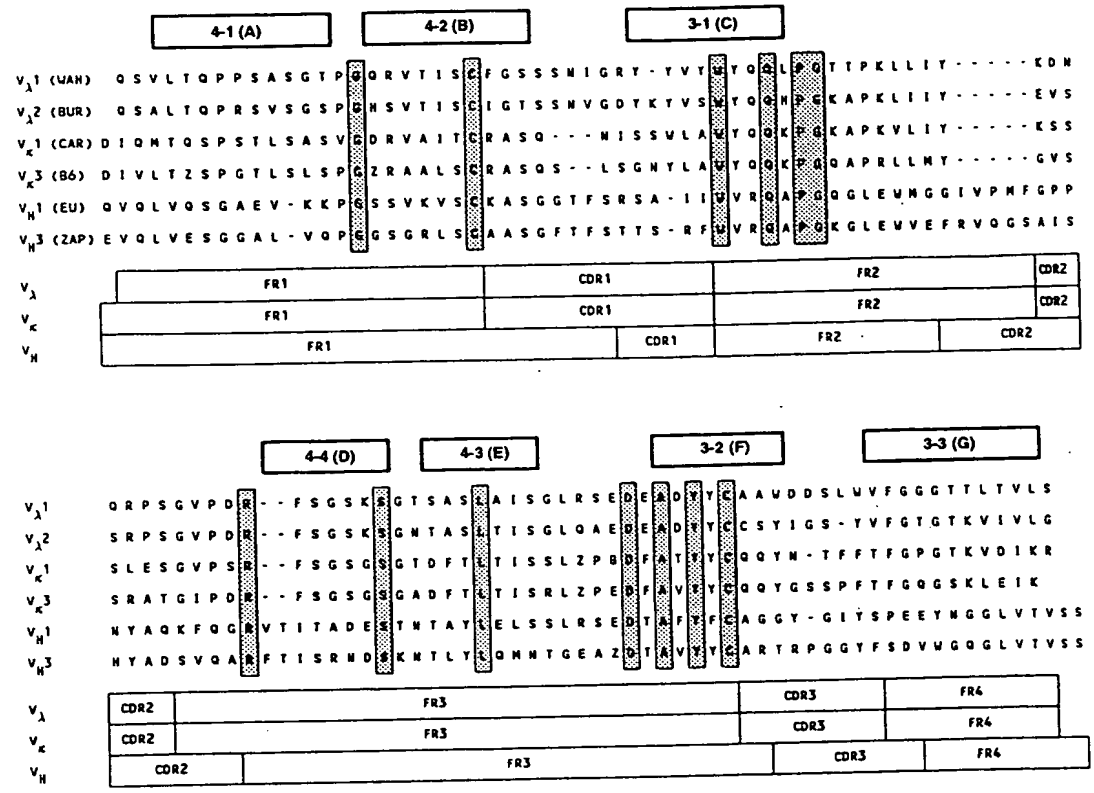


FIG. 5. Sequence alignment of six human V regions. Boxes above sequences represent β strands of the domains. Strands are numbered according to Edmundson and lettered (*in parentheses*) according to Hood. Gaps introduced to maximize homology are represented by *dashes*. Amino acids conserved among all six proteins are *boxed* and *shaded*. *Boxes* beneath the sequences depict the statistical boundaries of V region subdomains (see Table 1). Note differences in the lengths of the FR1, CDR1, FR2, and CDR2 segments between the four light chains and two heavy chains.

TABLE 1. *Boundaries delineating the statistical and structural subdomains of variable regions*

Subdomain region	Ig Chain	Residue positions	Boundaries of structural loop
FR1	Heavy	1–30	
	Light	1–23	
CDR1	Heavy	31–35*	H1: 26–32
	Light	24–34*	L1: 26–33
FR2	Heavy	36–49	
	Light	35–49	
CDR2	Heavy	50–65*	H2: 53–55
	Light	50–56	L2: 50–52
FR3	Heavy	66–94*	
	Light	57–88	
CDR3	Heavy	95–102*	H3: 96–101
	Light	89–97*	L3: 91–96
FR4	Heavy	103–113	
	Light	98–107*	

FR1–4 (framework regions), CDR1-3 (complementarity-determining regions), H1-3 (heavy chain variable loops), and L1-3 (light chain variable loops) are numbered according to Kabat et al. (1) using an alignment giving priority to conserved residues. Asterisks indicate regions which may have length variations depending upon germline V gene usage and/or junctional diversity. Data for the table were compiled from Kabat et al. (24a) and Chothia and Lesk (28a).

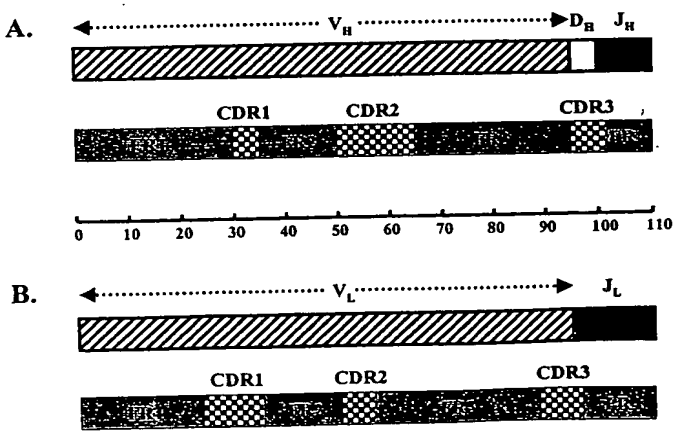


FIG. 6. Comparison of the gene structures and protein subdomains of rearranged V regions. (**A**) Heavy chain V domain. (**B**) Light chain V domain. The V_H and V_L gene segments are represented as *hatched boxes*, the D_H gene segment as a *white box*, and the J_H and J_L gene segments as *black boxes*. Framework regions (FRs) are displayed as *shaded boxes* and complementarity-determining regions (CDRs) are pictured as *checkered boxes*. An approximate amino acid scale separates the two diagrams.

development of V loci across several vertebrate species (see Fig. 7) for both heavy and light chains (40–44). Three V_H clans have been recognized using nucleotide sequence homology comparisons across the FR1 6–24 codon interval. While this stretch of FR1 sequence is conserved within a clan, a similar span (the 67–85 codon interval of FR3) can also be used to discriminate between V_H genes that belong to the same clan but differ in regard to the next level of classification, that of the family.

Classically, families are the groupings that have been used most frequently to identify and categorize V region genes relative to one another (reviewed in refs. 45–47). Members of a V region family share about 80% identity at the DNA level. Historically, when genomic Southern blotting was used to work out approximate family sizes, this degree of homology allowed for sufficient cross-hybridization to occur under low-stringency conditions, accounting for the utility of the family designation. At the protein level, this translates to levels of about 75% identity between gene products from the same family and less than 70% homology for proteins belonging to different families. Using these criteria, human V_H genes may be segregated into seven families, V_κ into four major families, and V_λ into ten families. The murine system (which contains larger absolute numbers of V genes) is more complicated, as evidenced by the fact that 14 V_H and 20 V_κ families have been recognized. Example sequences from several human V_κ and V_H gene families are aligned in Fig. 8. Note the presence of numerous "shared substitutions" within pairs of sequences belonging to the same family that are not present in the other sequences. These serve as distinctive "signature residues," which facilitate rapid identification of a particular sequence's "family of origin." The sequences in Fig. 8 also demonstrate another important characteristic of V gene families: Different families frequently possess different CDR lengths. Thus, independent of amino acid sequence, V region families intrinsically possess differing binding-site structures, thereby affecting their functional capabilities.

Shared substitutions have also been used to further refine families into subfamilies. Subfamilies, as are all classification schema to some extent, are particularly arbitrary divisions, largely because the parameters used to define subfamilies are not standardized. Generally, the features that describe a particular subfamily are esoteric and depend upon the specific characteristic(s) being studied. Finally, at the most descriptive level of classification, the work of Rabbitts, Winter, Honjo, and many others (46–50) has resulted in the identification, mapping, and—to some extent—sequencing of presumably most, if not all, human V region genes. Within a family single V gene segments may be compared against a consensus sequence representing that family (Fig. 9). When such a comparison is performed, it becomes clear that individual V genes' divergence is focused primarily in their CDR1 and CDR2 segments (42). Thus on the basis of primary sequence structural information, even closely related V genes may be predicted to adopt similar framework cores but differ in terms of their CDR1 and CDR2 loops such that a plethora of potential antigenic specificities are encoded genomically. In fact, even at the level of the individual V gene variability persists in this system. As one might expect with "variable" genes, allelic variation and polymorphism of the antibody locus exists as well—such that it is probably safe to conclude that all of the possible incarnations of immunoglobulin V genes will never truly be identified and categorized. Suffice to say, then, that the variable gene loci, clans, families, subfamilies, and even individual V genes themselves all derive from gene duplication events; only the period in evolutionary time at which the duplication occurred truly separates one gene or group of genes from any another.

While V regions provide the surfaces that interact with "foreign" antigenic determinants, constant regions perform the function of

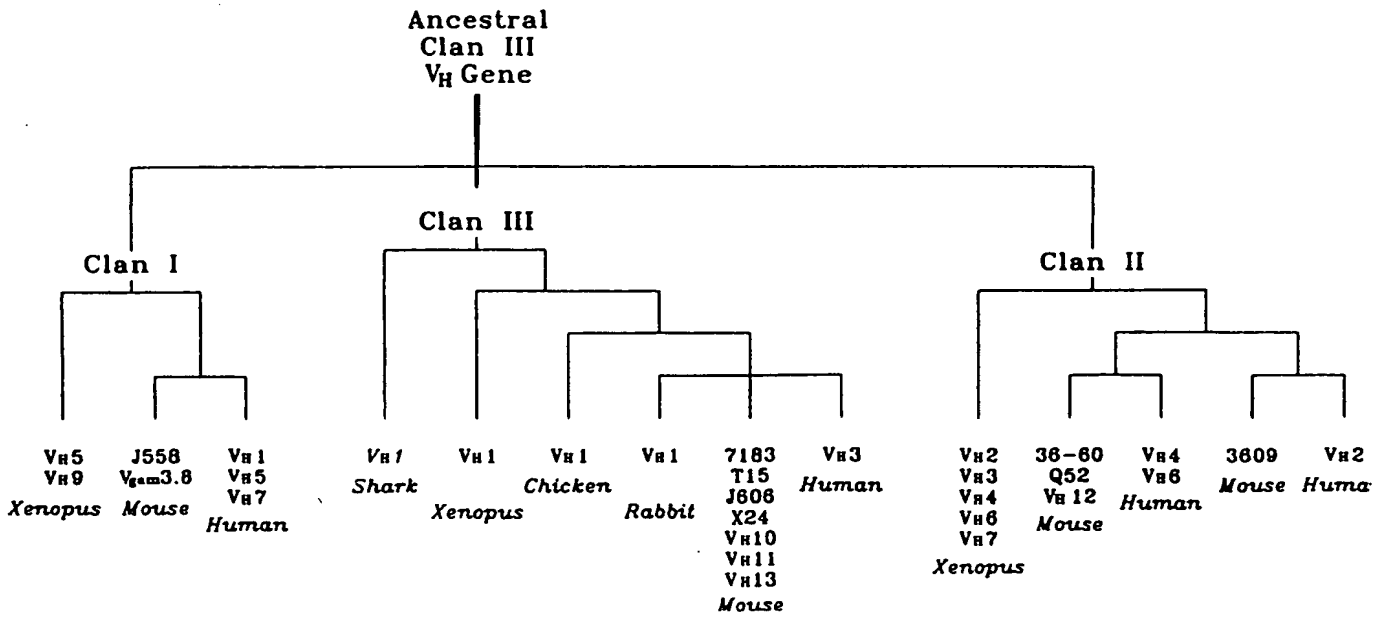


FIG. 7. Clan groupings segregate V_H genes across vertebrate species into distinct clusters of sequences. Note that several different V_H families in a given species can be present within a single clan, reflecting their underlying structural similarity. The lines depicting evolutionary relatedness are not drawn to scale. (From ref. 42, with permission).

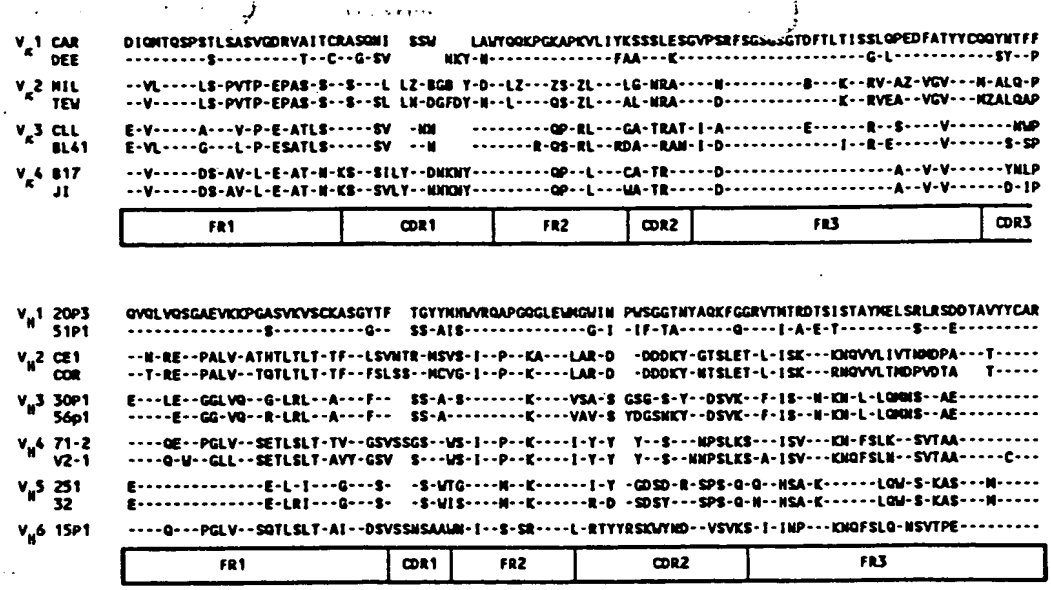


FIG. 8. Representative sequences from several human V_κ and V_H families. Gaps introduced to optimally align the sequences are indicated by *blank spaces.* Identity between residues in the top sequence and those below it is signified by *dashes.* Canonical boundaries for FRs and CDRs are schematized beneath both the light chain and heavy chain sequence groupings. Note the presence of many "shared substitutions" (amino acids common to two sequences belonging to the same family, but absent in sequences from other families).

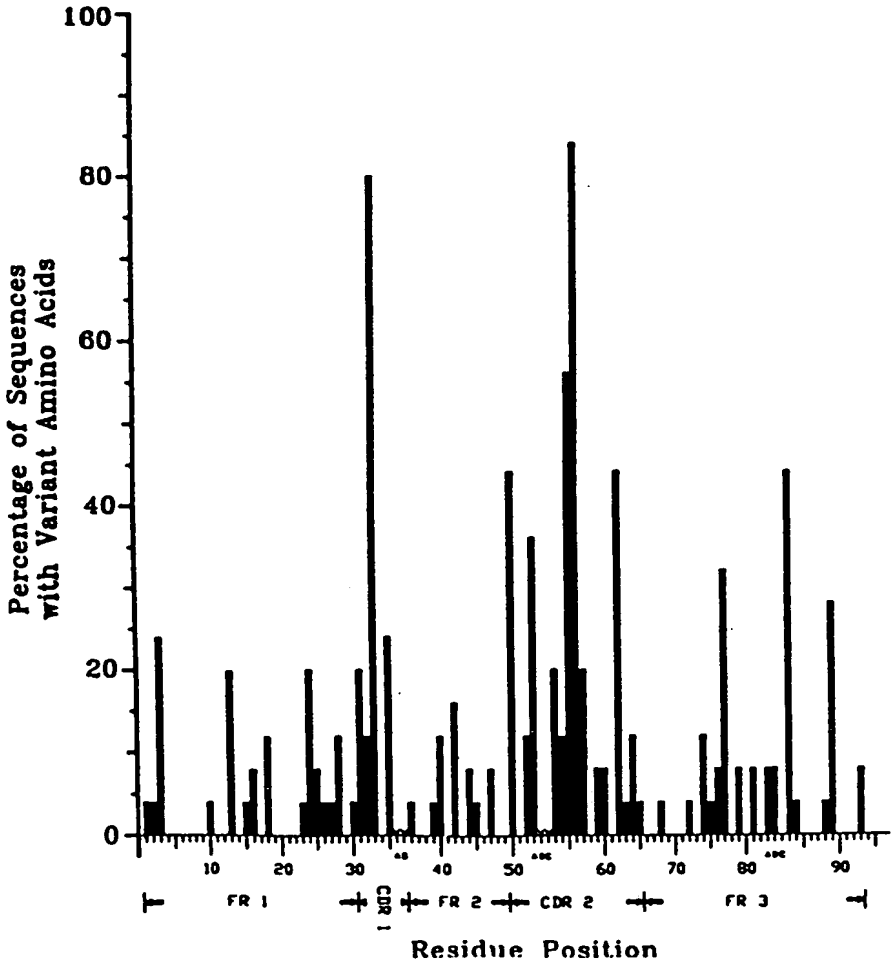


FIG. 9. Sequence variation within a V region family. Twenty-five unique germline sequences from the murine V_H 7183 family were used to generate a consensus sequence for the family. Divergence from the consensus 7183 sequence is represented as the percentage of sequences having a variant amino acid (an amino acid different from the consensus residue) at a particular position. Like the variability plot of heavy chains from multiple families (see Fig. 4), highly divergent positions within the same family also tend to localize to the CDRs. (From ref. 42, with permission.)

binding to the "self" molecules that mediate physiologic and immunologic effector pathways. Light chains have one constant region domain (either C_κ or C_λ) C-terminal to the V_L. The C_L pairs with the first constant region domain of the heavy chain (C_H1) using hydrophobic and disulfide interactions to stabilize heavy and light chain coupling. C_L is not known to specifically bind any other biological moities; therefore, no known effector properties are attributable to light chain C regions. By contrast, all known quaternary associations, biologic characteristics, and physiologic functions of immunoglobulins are governed by heavy chain C regions. Depending upon antibody isotype, heavy chain C regions consist of either three or four C_H domains. The first domain (C_H1) mediates association with C_L as detailed above and is part of the Fab fragment, while the following two (or three) C-terminal domains participate in interchain binding between heavy chain molecules and collectively comprise the Fc. Connection between C_H1 domains of μ and ε antibodies and their respective Fc is accomplished by specialized C_H domains ($C_\mu2$ or $C_\varepsilon2$) that allow for flexibility of the Fabs, yet maintain features of other Ig constant domains. Antibodies of the γ, α, and δ classes, however, utilize a shorter flexible hinge segment for this purpose, which will be specifically detailed elsewhere.

While initial studies indicated that C region sequences were reasonably highly conserved—at least relative to V regions (refer to Fig. 3)—their designation as "constant" is actually somewhat of a misnomer. For instance, studies comparing V and C regions from primitive species have suggested that the lengths of the loops connecting the β strands of constant regions actually vary more than do those of variable regions (51). Furthermore, between different isotypes of the same species, C_H regions share only about 30% sequence identity overall (Fig. 10). C region domains differ in regard to their interclass homologies such that different domains show various levels of sequence conservation. C_H1 domains are the most similar between isotypes; this may derive from the fact that all share the common function of pairing with immunoglobulin light chains. Similarly, the carboxy-terminal domains of μ, α, and γ ($C_\mu4$, $C_\alpha3$, and $C_\gamma3$) are substantially more related than the average for their Fc as a whole. As x-ray studies have indicated close contact exists between the C-terminal domains of the Fc, this sequence conservation may result from similar constraints as for the C_H1/C_L situation. Moreover, in the case of the μ and α chains, the relatively higher homology likely reflects the common role of the last domain in multimer formation. Also note that the $C_\mu4$ and $C_\alpha3$ domains each possess an 18-amino acid "tailpiece" at their C-termini. The penultimate cysteine residue in each sequence contributes to an intersubunit disulfide bond, allowing IgM and IgA polymerization (52). Much like the V region paradigm, a hierarchy of shared substitutions can be used to distinguish sequences of different domains, classes, and subclasses. In all comparisons between isotypes, however, it is apparent that the majority of conserved residues are localized to the β strands. As was the situation with the V region FRs, these strands are responsible for the proper folding of the domain. In addition, the two cysteines that form the intrachain disulfide bond, and the tryptophan which protects it from solvent reduction, are preserved among all C_H sequences as well.

In analogous fashion to V region evolution, the five heavy chain isotypes likely arose by duplication and diversification of a common gene precursor. This probably occurred in an ancestral organism that preceded mammalian speciation, because examples of all five classes appear in all mammals. Actually, interspecies sequence comparisons of a given isotype demonstrate their similarity far exceeds that of the different isotypes within a species (compare

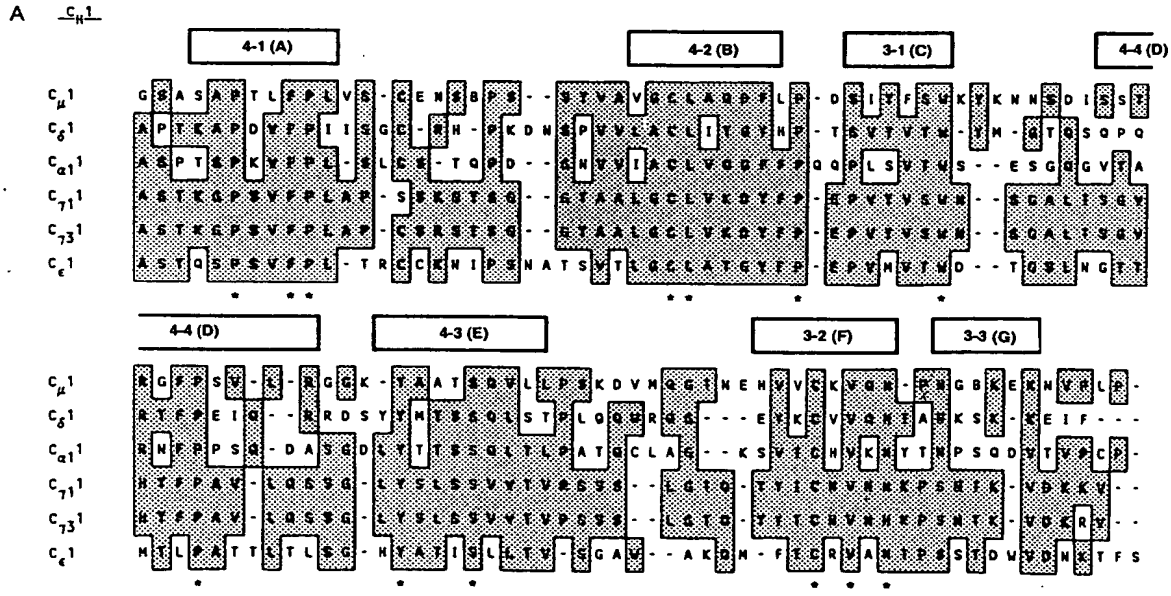


FIG. 10. Comparison of the amino acid sequences of the **(A)** C_H1 and **(B)** Fc regions of all human isotypes (excluding the $\gamma2$, $\gamma4$, and $\alpha2$ subclasses). *Boxed* and *shaded* amino acids represent residues shared by two or more isotypes. *Asterisks* mark amino acids conserved among all six sequences. *Dashes* indicate gaps introduced to maximize homology between sequences. β strands are numbered according to Edmundson and lettered (*in parentheses*) according to Hood in *white boxes* above the alignments.

Figs. 10A and 11). Subclasses represent more recent gene duplication events. Accordingly, distinct subclass profiles exist in various mammalian species. For instance, while humans have two α loci and hence two IgA subclasses (53), murine species have only one α gene, and rabbits have 13 (54)! Due to their later evolutionary divergence, heavy chain subclasses display greater sequence similarity—on the order of 60% to 90%—as evidenced by the γ1 and γ3 alignments depicted in Fig. 10. Despite the high levels of concordance between C regions of related subclasses, even slight differences can have profound functional repercussions. As an example, among the four human IgG subclasses (whose sequences are over 95% identical to one another), IgG1 and IgG3 bind to macrophages and other phagocytes with ease, while IgG2 and IgG4 bind very poorly. This binding is mediated by an Fcγ receptor that has been extensively characterized (detailed later within this chapter). Similar subtle sequence differences are also involved in functional properties affecting immunoglobulin catabolism, placental transfer, and reactivity with staphylococcal protein A (SPA). It is this selective pressure for the ability to perform in a variety of functional capacities that has both maintained the five major C_H classes throughout mammalian species and also driven the evolution of subclasses to their points of divergence in these same

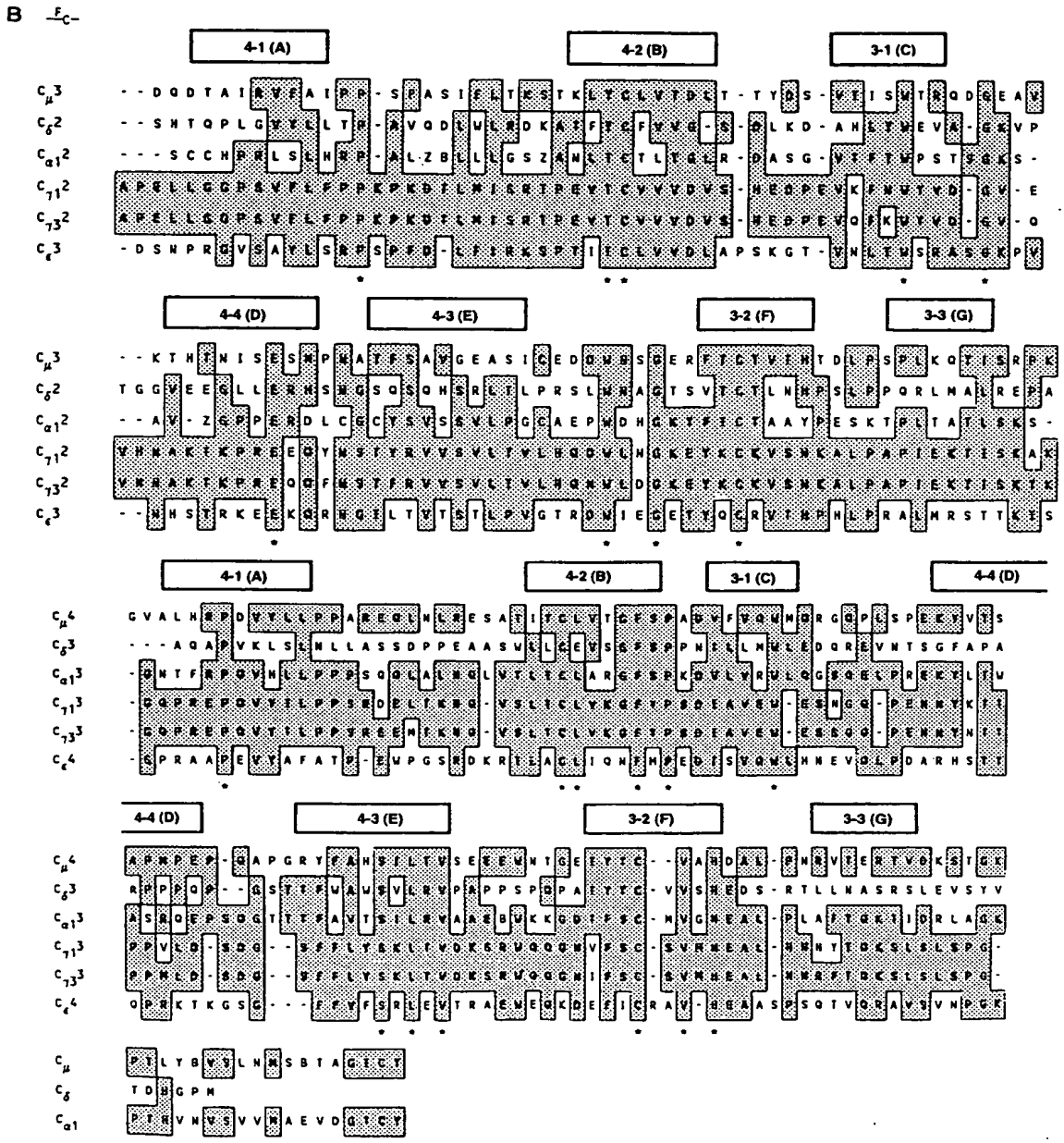


FIG. 10. *Continued*

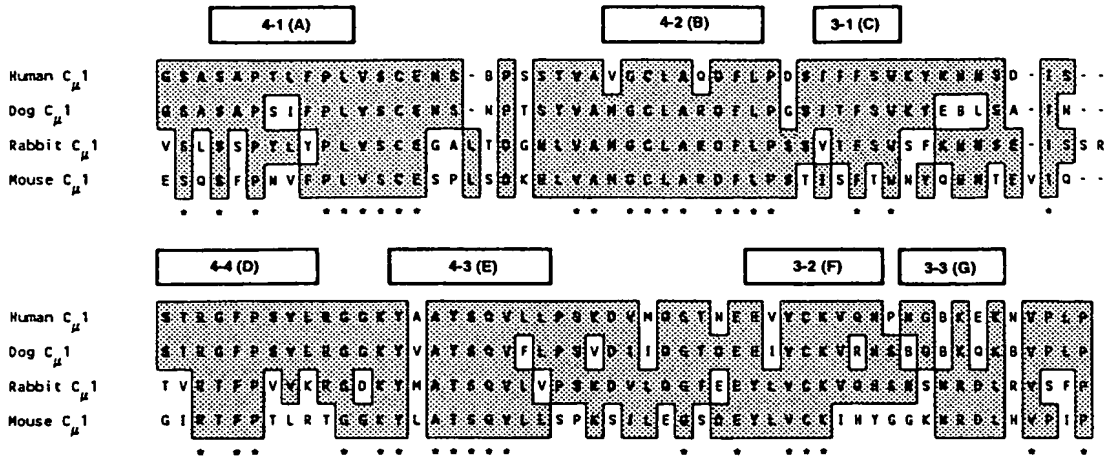


FIG. 11. Protein sequence alignments of the $C_\mu 1$ domains from four mammalian species. Labels are as for Fig. 10. Note the paucity of gapping required and also the high number of absolutely conserved residues (*) across species relative to the alignments of the different human $C_H 1$ isotypes in Fig. 10A.

species. Collectively, these evolutionary changes have bestowed upon the different isotypes the ability to respond to antigenic challenge in a variety of immunologically productive ways.

Despite the absence of evidence to suggest that light chain C regions participate in biologically significant interactions (other than coupling with heavy chains), multiple forms of C_L have also been identified. First, the light chain isotypes C_κ and C_λ—while functionally indistinguishable—exist as separate genetic loci with their own complement of possible V and J gene segment partners. This is in contrast to the situation with the heavy chain locus, in which any rearranged VDJ can potentially become associated with any C_H isotype via the process of class switching (see Chapter 24). While the use of κ versus λ isotypes seems to be inconsequential as to the antibody's efficacy, differences in their utilization are nonetheless present. In human immunoglobulins the κ:λ ratio is approximately 70:30, while in murine systems about 95% of antibody is of the κ class (20). The reasons for these imbalances are yet to be definitively explained, but are probably related to the number of V_κ and V_λ genes available for use in the respective genomes of these organisms. As was the case for heavy chain C regions, light chain isotypes are well conserved across species (Fig. 12). In addition, within a species, κ and λ classes are more similar to each other (about 38%) than were the different C_H isotypes (compare Figs. 10 and 12). Note also the presence of a terminal (C_κ) or penultimate (C_λ) cysteine residue in these sequences. This half-cystine is responsible for the light chain's contribution to the H–L interchain disulfide bond. A second point of variation within C_L regions concerns C_λ subclasses. While only a single C_κ gene is present in human and mouse, five or four λ-chain subclasses exist in human or murine genomes, respectively. A final deviation from "constancy" involves the C_κ domain. While only one κ constant region gene is found in the locus, three human κ allotypes have been identified. These alleles differ in regard to their residues at the surface-exposed positions 153 and 191, such that these allotypic markers are able to serve as antigenic determinants (55).

As mentioned previously, hinges—either as distinct domains in the cases of the μ and ε chains or as shorter specialized segments for α, γ, and δ antibodies—connect the Fab and Fc portions of the immunoglobulin molecule. However, the lack of extensive hinge structure in nonmammalian heavy chain sequences indicates that the evolution of this function largely occurred after mammalian radiation (51). Consequently, discussion here is restricted to mammalian hinges as typified by human sequences.

Hinge regions not only connect Fab to Fc, but also contain the disulfide bonds that covalently link the two heavy chains (discussed further in the section on quaternary structure) in the Fc portion of the antibody molecule. They display great variability between isotypes and are generally encoded by unique exons. For example, the hinges of human IgG1, IgG2, and IgG4 are each produced from a single short exon encoding a peptide of between 12 and 15 amino acids. Alternatively, the IgG3 hinge derives from four distinct exons, resulting in a hinge region that spans approximately four times as many residues as the other γ isotypes (56). Owing to similarities between the different hinge sequences and the extra C_μ and C_ε domain sequences, it is thought that hinges evolved from the $C_{\mu/\varepsilon}2$ domain; unfortunately the hinge sequences are too short and the homologies too weak to trace hinge lineage with certainty. It is clear from sequence comparisons, however, that the $C_\gamma 3$ hinge evolved from the $C_\gamma 1$ hinge by multiple duplication (further substantiated by the aforementioned $C_\gamma 1$ and $C_\gamma 3$ hinge exon arrangements). Figure 13 aligns the extra domains of C_μ and C_ε, in addition to the hinges of several other human isotypes.

In addition to the variable, constant, and hinge regions of antibodies, other primary sequence features must also be considered. For instance, both heavy and light chains of immunoglobulin are synthesized with a leader peptide (almost entirely encoded by a separate exon upstream of the V region) of 16 to 26 amino acids in length. During translation, this leader (more generally referred to as a signal peptide) directs the mRNA/ribosome/polypeptide complex to the rough endoplasmic reticulum (RER) where translation is completed. During the synthesis and extrusion of the nascent polypeptide through the RER membrane, the signal peptide is then removed by specific proteolytic cleavage.

Finally, for a comprehensive discussion of primary immunoglobulin structure, mention must be made of sequences present exclusively in the case of surface immunoglobulin expression. Heavy

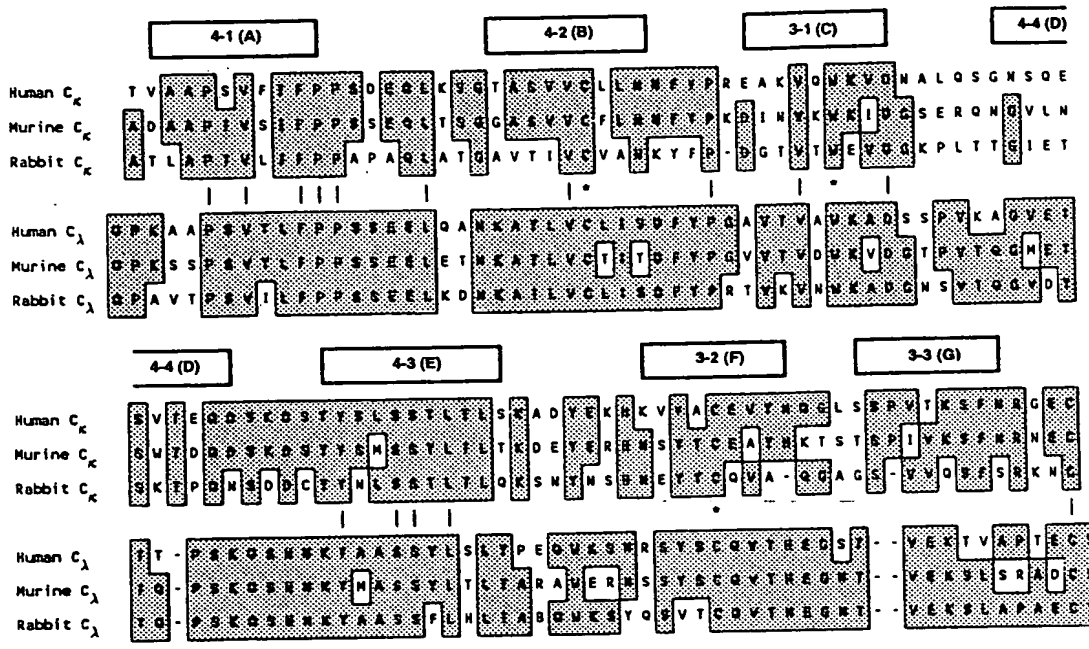


FIG. 12. Comparison of the C_κ and C_λ domains from three mammalian species. Labeling is as for Figs. 10 and 11, except *vertical lines* now represent amino acids shared among all six sequences and *asterisks* specifically denote the invariant tryptophan and cysteines at the core of the domain (also shared by all sequences). As in Fig. 11, little gapping is needed and high homology persists across species.

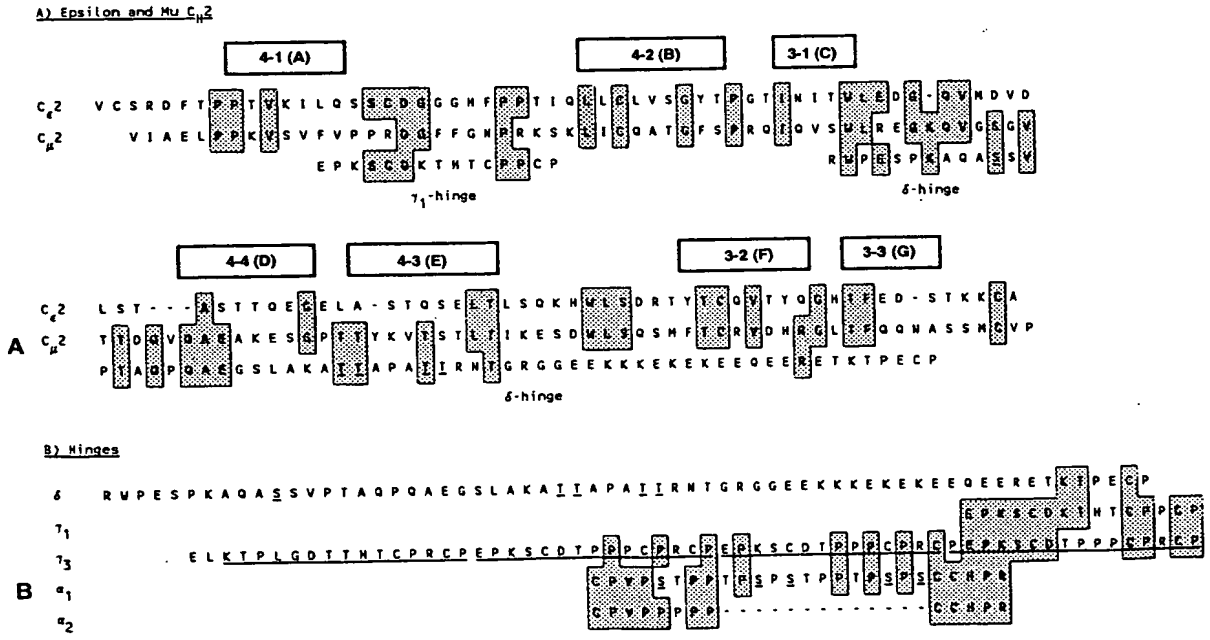


FIG. 13. Alignment of the amino acid sequences of **(A)** the human $C_\mu 2$ and $C_\epsilon 2$ domains and the γ1 and δ hinges and **(B)** several human immunoglobulin hinge regions. Labeling is as for Figs. 10–12. In **(A)**, the $C_{\mu/\epsilon}2$ domains are compared with the hinges of two different isotypes to display their potential, although limited, evolutionary relationship. Part **(B)** demonstrates several features of hinge regions. Note the high proline, cysteine, and serine/threonine content of the hinges, which consists of the bulk of the homologies between isotypes. This interclass homology is reasonably low, whereas intersubclass homologies (compare γ1 to γ3 and α1 to α2) are considerably higher. Note the fourfold duplication of the γ1 hinge in γ3 (the repeating unit is *underlined*). Also note the deletion within the α2 hinge relative to α1. O-linked glycosylation sites in the δ and α1 hinges are also *underlined*.

chains of all isotypes can exist either as secreted antibody or as membrane-bound immunoglobulin (mIg), which serves as the central component of the antigen-specific B cell receptor (BCR). The choice between mIg and secreted immunoglobulin is manifest at the level of alternative mRNA splicing at the 3′ end of the message. Differential processing in favor of the surface immunoglobulin form results in the replacement of the hydrophilic carboxy-terminal residues of secreted antibody with a stretch of hydrophobic amino acids (which anchors the immunoglobulin in the cell membrane) and a short cytoplasmic tail (57,58). Expression of these sequences not only tethers the immunoglobulin to the cell surface, but also governs the ability of mIg to interact with other constituents of the BCR necessary for signal propagation and eventual activation by antigen (see Chapter 7). As such, the regulation of this splicing event and its ultimate protein products is tightly controlled through the stages of B lymphocyte differentiation so as to ensure the proper production of membrane immunoglobulin versus secreted antibody at the appropriate developmental stage of the B cell (see Chapters 5 and 6).

Secondary Structure—The Immunoglobulin Fold

With the exceptions of the hinge and cytoplasmic tail, the properties of the immunoglobulin fold as a protein motif dominate all aspects of immunoglobulin structure, from primary to quaternary. In regard to secondary structure, this refers to the different patterns of β-pleated sheet formation assumed by the V and C regions. As explained earlier, all immunoglobulin folds are composed of two layers of antiparallel β sheet arranged in a sandwich (or β-barrel) that encloses a hydrophobic interior. Early on, it was recognized that each immunoglobulin domain contains seven polypeptide β strands, four of which comprise one β-pleated sheet, the other sheet consisting of the remaining three strands. Accordingly, a nomenclature was developed that reflected within which β sheet (four-stranded or three-stranded) a particular strand was located, using numbering (3-1, 3-2, 3-3 and 4-1, 4-2, 4-3, 4-4). Subsequent studies revealed that superimposed upon this secondary structural organization shared by all immunoglobulin folds was a discrepancy between V and C region immunoglobulin folds; some of the "extra" amino acid residues found in variable regions actually participated in β sheet formation, giving rise to two additional β strands. Consequently, the immunoglobulin folds of V domains actually form a barrel using five-stranded (analogous to the C region three-strand sheet) and four-stranded β-pleated sheet layers. A second naming system using letter designations (A, B, C, C′, C″, D, E, F, and G) for the different strands of immunoglobulin fold β pleated sheets makes allowance for these extra β strands. Figure 1 displays "unfolded" immunoglobulin folds of each type so as to schematically represent the secondary structure of both V and C immunoglobulin domains. Note that while the letter nomenclature proceeds in accordance with the primary structure of the protein (A, B, C, D, etc.), the numbering system of β strands is nonlinear with respect to the primary sequence (4-1, 4-2, 3-1, 4-4, 4-3, 3-2, 3-3). Rather, this naming system is designed so as to coincide with the three-dimensional orientation of the strands in the context of the fully folded immunoglobulin domain (refer back to Fig. 2), which will be discussed in greater detail below.

Tertiary Structure—The Immunoglobulin Domain

Logically, if the β sheets of the immunoglobulin fold are the dominant secondary structural protein motif of the antibody, then the fully-folded tertiary structural correlate of this paradigm is the of the immunoglobulin domain. As has been explained, Ig domains are founded upon the premise of two layers of β-pleated sheets sandwiched around a core of hydrophobic side chains, to form compact globular structure. Of course, upon this general structural framework, Ig domains vary considerably. Still, certain features common to all Ig domains of actual immunoglobulin molecules (as opposed to Ig domains of other IgSF members) are invariant. Central to the Ig domain, both literally and figuratively, is the presence of the two cysteines that form the intradomain disulfide bond an

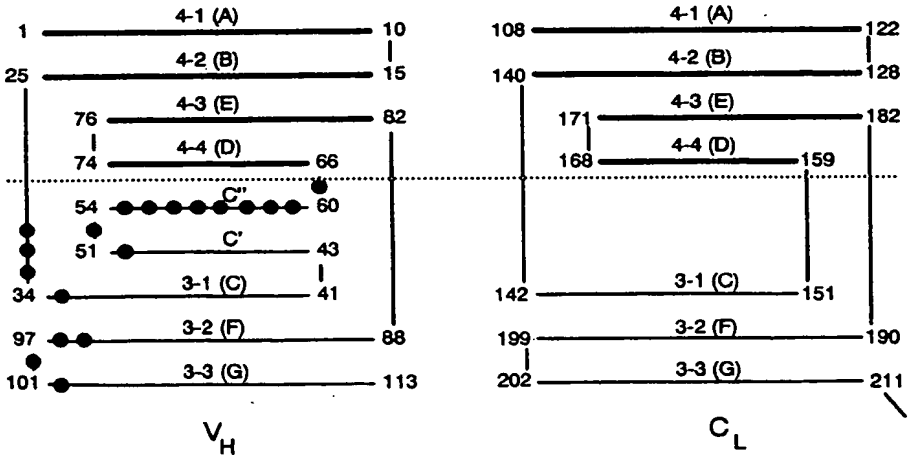


FIG. 14. Schematic diagram of the secondary structure of "unfolded" immunoglobulin domains. β strands are represented by *horizontal lines*, loops connecting β strands are depicted as *vertical lines*. *Bold horizontal lines* are in the four-stranded face of the domain, *light horizontal lines* are in the three- (five-) stranded β sheet. *Dotted regions* denote CDRs. Strands are numbered according to Edmundson and lettered (*in parenthesis*) according to Hood. In three dimensions, immunoglobulin domains are folded with the *light horizontal strands* under the *bold horizontal strands* using the dotted line as an axis.

the tryptophan which protects this bond from hydrolysis. All domains of immunoglobulin—variable region or constant, heavy chain or light–conserve these three key residues, and they occupy homologous positions in the different domains of the fully-folded protein.

Beyond these key core residues, however, different Ig domains are still able to maintain similar tertiary structures in the face of dramatic differences in their primary sequences. This chiefly derives from the fact that the identity of a particular residue at a particular position is not nearly so important to proper Ig domain folding as is the character of the particular residue at that position. In other words, as long as amino acids with side chains compatible with β-pleated sheet formation are present in the proper locations and those necessary to terminate β strands are similarly in the correct places, their actual identity appears not to be crucial.

There are, of course, other residues essential to proper folding and functioning of Ig domains, but these are specific to particular domains and not common to all Ig domains within immunoglobulin. For example, the FR4 motif W/F-G-X-G is widely conserved, but only among V regions, where it serves to create a "β bulge" necessary for proper V_H/V_L dimerization. Similarly, the V_H-specific (G-L-E-W-hydrophobic) and V_L-specific (P-hydrophilic-hydrophobic-L-hydrophobic) FR2 sequence motifs—and their accompanying β bulges—are also conserved tertiary structures among their respective subsets of Ig domains. Finally, as was discussed in secondary structure, V and C domains also differ in terms of their basic arrangement of the immunoglobulin fold such that V domains are composed of a four-strand (A, B, E, and D) and a five-strand (C", C', C, F, and G) layer, whereas C domains consist of four-stranded (A, B, E, and D) and three-stranded (C, F, and G) β sheets. Clearly, this distinguishes variable and constant domains at the tertiary structural level as well.

Once again, as with primary and secondary structural considerations, examination of tertiary structure can be facilitated by distinguishing between V and C domains. In the case of variable regions, whether from heavy or light chains, obviously no single protein structure will ever fully suffice in describing the entire group. This is because each domain, V_H or V_L, is in effect a new structure, and unless solved crystallographically, can only be postulated. Nonetheless, two broad generalities as pertains to V region tertiary structure are evident. First, the similarity between two different V domain structures tends to closely parallel their relatedness at the genetic level; that is, one can reasonably predict that two different V_H structures will be more similar to each other than to a V_L structure, two V_H domains belonging to the same clan will be more similar than to one from a different clan, and so on. Colorplates 1 and 2, which compare molecular models of FR1 regions from antibodies belonging to the same and different clans, are particularly compelling in this regard. While undoubtedly exceptions exist where two proteins are genetically similar but diverge at a few key residues with important structural consequences, in most instances this rule is valid for extrapolating the tertiary structures of unsolved V domains. In this context, it is perhaps also important to note that the FR1 region utilized to assign clan identity is solvent exposed and distal to the antigen-binding site, while the FR3 region which correlates best with family designation is immediately adjacent to the binding site, and capable of affecting its conformation and even interacting with antigen directly (42). In light of several reports linking over-representation of certain families in the repertoires reactive against particular antigenic specificities, the tertiary structural predictions available by this means perhaps take on added significance.

The second pervasive trend which is apparent upon scrutiny of V region tertiary structures is the tendency of the antigen-binding site to represent a "nested gradient of antibody diversity" (42). Recall, investigations into the very first antibody proteins identified nonconserved "hypervariable" regions and reasonably well-conserved "framework" regions, which were in turn postulated to correspond to the antigen-binding and structural foundations, respectively, of the molecule. These hypotheses proved correct, as the FRs were demonstrated to coincide with the β strands, and the CDRs were shown to derive from the variable loops that interconnect the strands. Moreover, current understanding of antibody structure makes it possible to recognize that, as a general rule, the most variable residues of an immunoglobulin V region localize immediately proximal to the antigen-binding site, whereas those that are most conserved tend to be distant to that site. Colorplate 3 provides one such example of this concept. Thus, the three-dimensional context in which amino acids interact to create a platform for ligand-binding (paratope) diverges dramatically from antibody to antibody.

There are several factors that influence the tertiary structure of the paratope itself, and their composite effect can be complex. First, sequence variation of two types in the loops can profoundly alter antigenic specificity and affinity. CDRs vary considerably in length, both as a function of V gene usage (affecting primarily CDRs 1 and 2) and as a consequence of junctional diversity (affecting only CDR3). Second, CDRs obviously differ significantly in terms of their sequence composition, due once again to gene usage and junctional diversity, and in addition to somatic hypermutation. In this way, diversity-generating mechanisms target amino acid variability to the CDR loops where they are most apt to change both the physical shape and chemical nature of the combining site. Also, because FR residues near CDR boundaries frequently interact with antigen directly (59), alterations in these positions affect structural variety as well. Even glycosylation of CDR asparagine residues has been implicated in changing loop conformation and antigen binding (60,61). Conformational variability can also play a critical role in diversifying the paratope surface, as CDR loops have been shown to interact extensively with adjacent FR amino acids and with each other (62,63). These studies, and failed attempts to engineer antibodies by simply swapping CDRs onto different FR backbones without appreciably affecting affinity for antigen, have further demonstrated that while V regions can be conveniently dissected into FRs and CDRs at the primary structural level, in actuality these elements cooperate to facilitate antigen-binding rather than acting as discrete elements.

While the limitless capacity of the immune system to generate new variable region domains makes impossible the absolute elucidation of all potential V region structures, the relatively smaller number of constant regions has allowed reasonable progress to be made in terms of assigning definitive tertiary structures to the C region domains. To date, x-ray diffraction analysis has resulted in a high resolution structure for only the Fc fragment of IgG1 (11). An α-carbon backbone of this structure is shown in Fig. 15A. Additionally, one whole immunoglobulin (an unusual IgG1 molecule with a hinge deletion) has been crystallized (64). Other Fc isotypes, based on the Fcγ1 paradigm, can be modeled (see Fig. 15B) using sequence homology with IgG and energy minimization calculations (65). Otherwise, almost all three-dimensional studies have utilized Fab fragments or Fv fragments, often produced in

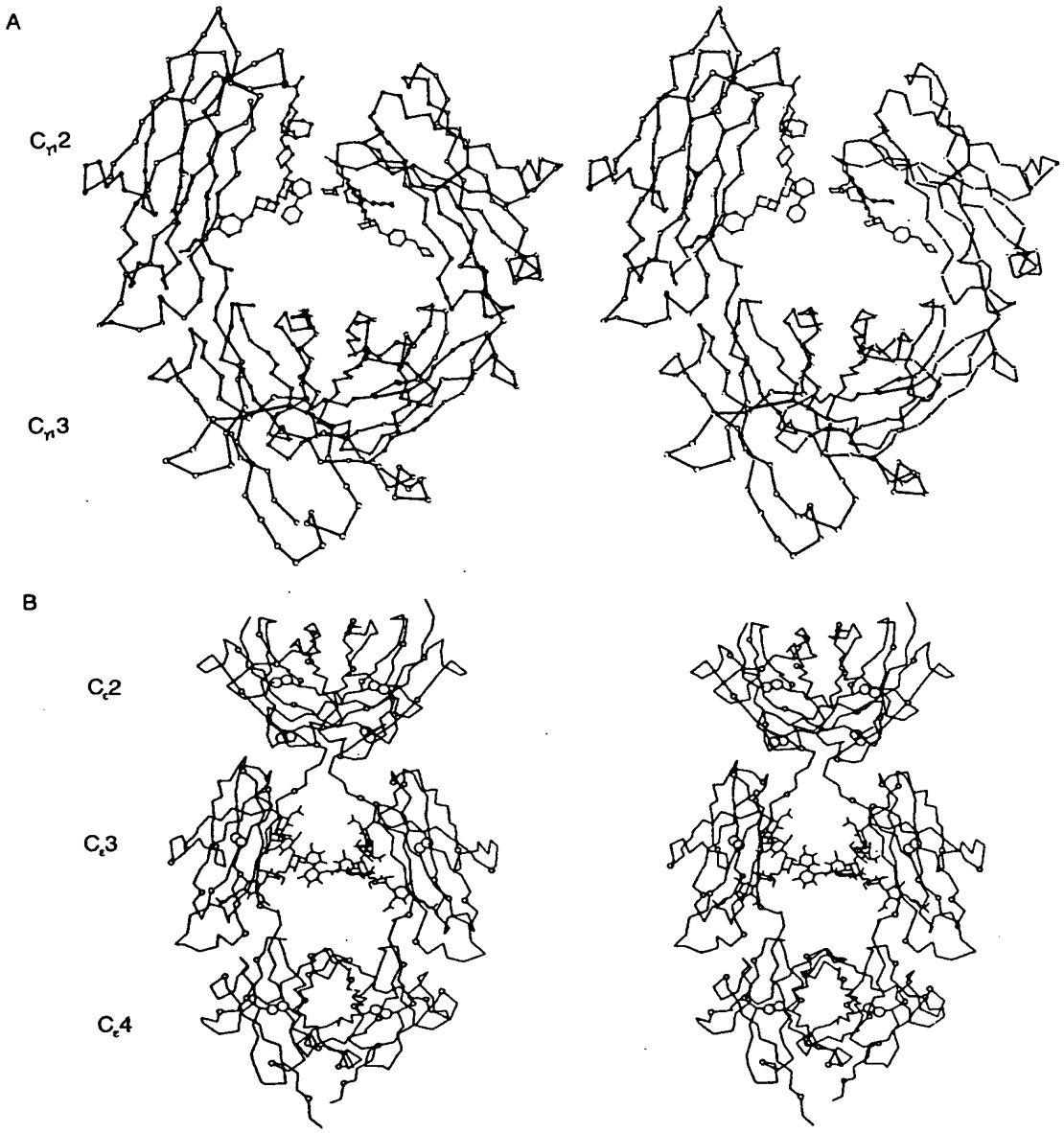
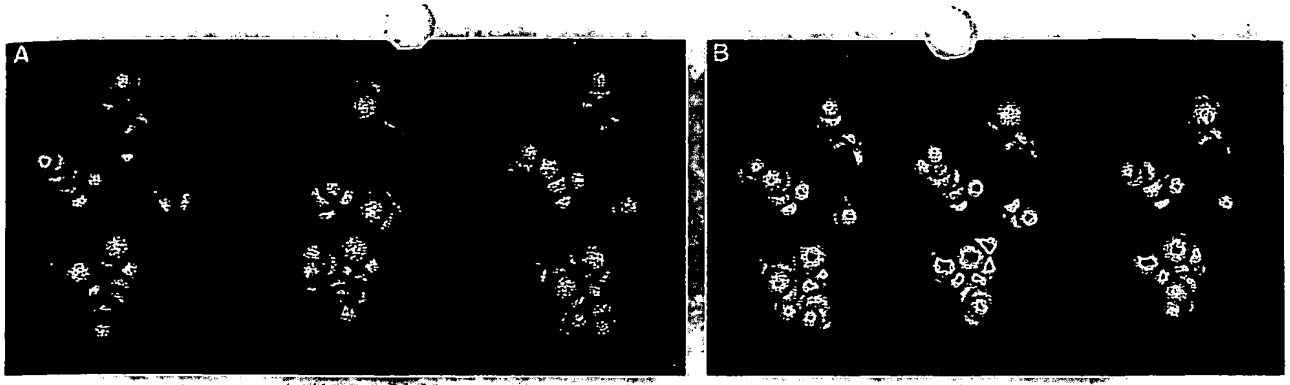


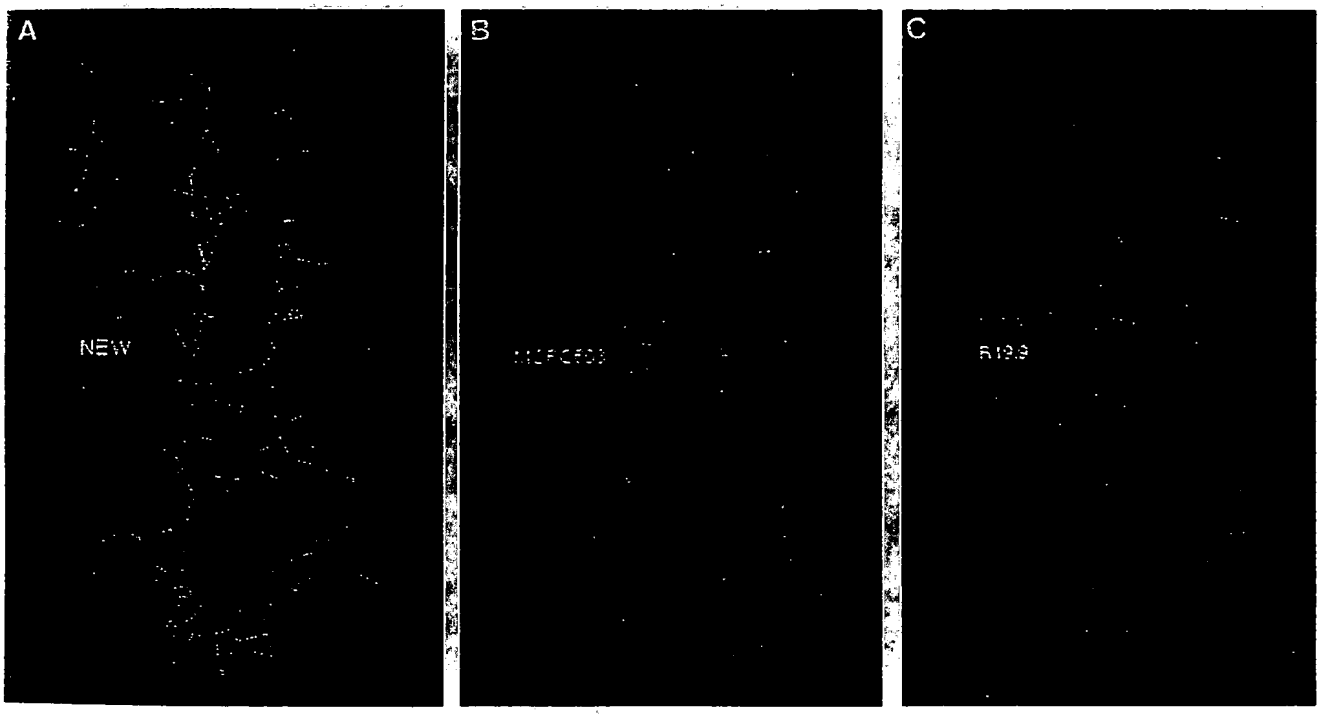
FIG. 15. Stereoviews of the α-carbon backbones of (**A**) Fcγ1 and (**B**) Fcε. In both cases, the C-terminal domains dimerize, but the penultimate domains ($C_\gamma 2$ or $C_\varepsilon 3$) contact their counterparts only via their carbohydrate moieties. The Fcε structure is a prediction modeled on the IgG1 crystal. As IgG has no $C_\varepsilon 2$ domain equivalent, the $C_\varepsilon 2$ structure is less reliable conjecture than is the rest of the molecule. (From refs. 12 and 65, with permission.)

bacteria. Of note, the IgG Fc crystal has also been solved as a co-crystal with staphylococcal protein A (SPA) (12). Importantly, this structure reveals the binding of SPA between both the $C_\gamma 2$ and $C_\gamma 3$ domains. This is contrary to the notion originally promulgated that single domains would perform IgG functions. Another example of sharing of function between domains is the Fcα receptor binding site on human IgA, which also involves both $C_\alpha 2$ and $C_\alpha 3$ (66). Surprising results such as these demonstrate the need for continuing investigation—by both crystallographic and other means—into the structural intricacies of all of the immunoglobulin isotypes. In any case, in the intervening time, the general properties of topology, though proven only for IgG, are reasonably generaliz to the other isotypes, since the residues involved in the various cc tacts are largely conserved.

As for the V region domains, constant region domain structure governed by general principles that tend to hold true in most exa ples of C region domains thus far studied. The overriding cc sideration in this regard, as for the V domains, concerns the re tive concentration of variability outside the β strands. C regic conserve a much higher percentage of residues from domain

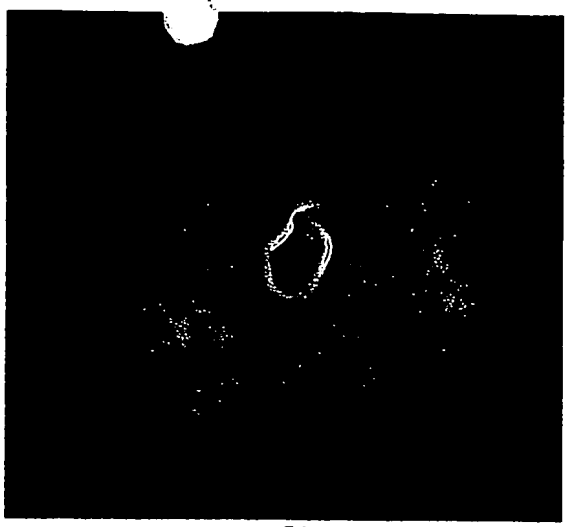


COLORPLATE 1. Three-dimensional models of immunoglobulin FR1 regions. Space-filling representations of amino acids 6–24 are displayed with orientation such that the CDRs would be above, and the C_H1 domain below, the models. Amino acid residues 6, 9, 12, 13, 16, 19, and 23 are colored *yellow* for reference. In (**A**), antibodies deriving from each of the three clans—clan I HyHEL-5 (*left*), clan II NEW (*middle*), and clan III KOL (*right*)—are presented. Clearly these structures differ in their structural characteristics. In (**B**), three clan III antibodies—human V_H3 family KOL (*left*), murine V_HT15 family MCPC603 (*middle*), and murine V_HX24 family J539 (right)—are compared. Note the similarity of these three FR1 loops relative to those in part (**A**). (From ref. 41, with permission.)



COLORPLATE 2. Superposition of FR1 regions from antibodies of (**A**) different, and (**B,C**) the same, clans. Stick diagrams of amino acid residues and their respective side chains are overlayed to facilitate comparisons. In (**A**), clan I HyHEL-5 and clan II NEW are superimposed on clan III KOL. These molecules can be seen to vary significantly. In (**B**), clan III antibodies MCPC603, J539, and KOL are modeled. Note that the agreement between structures even extends to side-chain sizes and orientations. In (**C**), clan I R19.9 antibody is superimposed on clan I HyHEL-5 with similar results. (From ref. 41, with permission.)

(a)



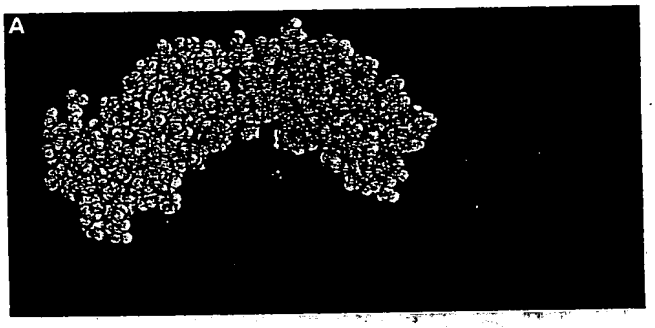
(b)

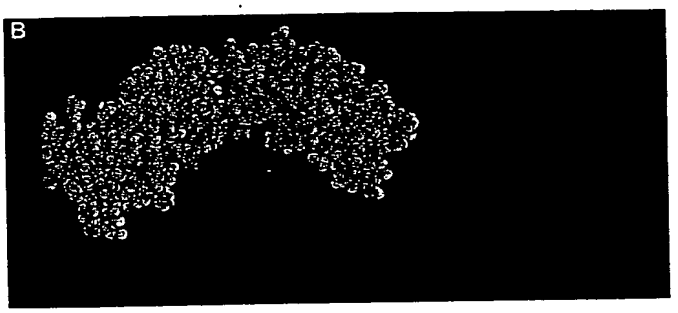
(c)

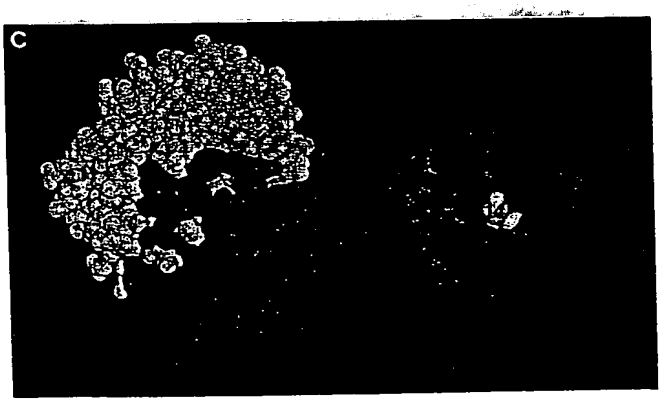
(d)

COLORPLATE 3. The antigen-combining site is the product of a nested gradient of diversity. Antibody sequence variation is mapped on the template of the surface of the antibody POT using a scale of *blue* (highly conserved) to *red* (highly divergent). The V_κ domain is to the left of each model, and the V_H domain is on the right. The most highly variable CDR3 loop of the V_H is depicted as a *gray ribbon* in the center of the diagrams; the highly variable V_κ CDR3 is not shown in these representations. In (**A**), germ-line diversity is displayed. In (**B**), diversity introduced by somatic hypermutation is presented. In (**C**), the sum of these two diversities is depicted. Note that in all cases, variation is predominantly restricted to the antigen-binding site, and diversity is highest in close proximity to the V_H CDR3. In (**D**), residues that have been demonstrated to make direct side-chain contacts with antigen in 21 separate crystals are plotted on the same *blue* (zero contacts) to *red* (up to 16 contacts) scale. Thus, the presence of diversity and the tendency to contact antigen are intimately related. (From ref. 58a, with permission.)

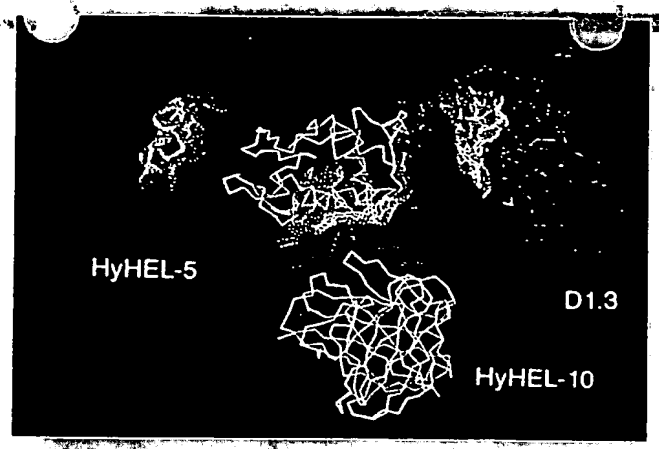
COLORPLATE 4. Ribbon diagram of the α-carbon backbone of an immunoglobulin Fv fragment. The heavy chain is shown in *blue*, and the light chain is represnted in *violet*. The invariant cysteines, tryptophans, and aromatic residues at the core of the V_H and V_L domains are shown in *yellow*. Orientation of the Fv is such that the antigen-binding site is at the *top*, and the C_H1/C_L domains would lie at the *bottom*, of the plate. (From ref. 42, with permission.)



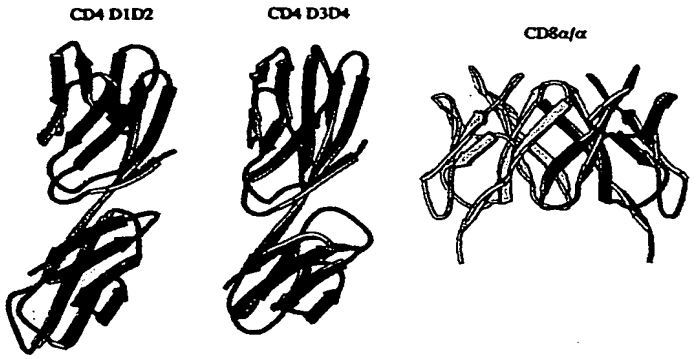




COLORPLATE 5. Model of the lysozyme–antilysozyme antibo (Fab D1.3) complex. The lysozyme molecule is depicted in *gre* and its glutamine-121 residue in *red*. The D1.3 heavy chain is sho in *blue*, and the light chain in *yellow*. In (**A**), the complex is seen it was in the crystal—as a blunt, end-to-end interface of the two m ecules. In (**B**), the two proteins have been separated to demonstr their structural complementarity. In (**C**), the molecules have be rotated toward the viewer to allow visualization of important cont amino acids (now portrayed in *red*, Gln-121 in *violet*). (From 109a, with permission.)



COLORPLATE 6. Model of lysozyme and three antilysozyme complexes. The Fab D1.3 (see Colorplate 5) is included for reference along with two additional antibodies: HyHEL-5 (*left*) and HyHEL-10 (*below*) Molecules are shown as α-carbon backbones except for colored van der Waals surfaces involved in binding. (From ref. 110 with permission.)



COLORPLATE 7. Ribbon diagrams of CD4 D1D2 (*red*), CD4 D3D4 (*green*), and a CD8α/α homodimer (one subunit *yellow*, the other *blue*). Note the continuous β strand that links domains D1 to D2 and D3 to D4 in CD4. This causes the D1D2 and D3D4 segments to be rigid structurally. (From ref. 207, with permission.)

domain—across immunoglobulin class and even across species—than do individual V regions (refer back to Figs. 10–12), but even so, stretches of relatively lower identity can be localized to distinct parts of the protein. Even though C domains are not formally delineated into framework and hypervariable regions, the tenets used to classify the subdomains of V regions are still applicable when discussing C regions. In Fig. 10A, for example, which compares different human C_H1 domains, note the areas of not only lower conservation but also where gapping is necessary to preserve alignment. Without exception, these regions are most prevalent between strands, especially the loops connecting strands 4-1 and 4-2 (A and B), 4-2 and 3-1 (B and C), 3-1 and 4-4 (C and D), 4-4 and 4-3 (D and E), and 4-3 and 3-2 (E and F). Conversely, the areas of highest conservation are found within the β strands (especially 4-1, 4-2, 3-1, 4-3, and 3-2), where, in addition to the residues needed for the intradomain disulfide linkage present in all Ig domains, the amino acids necessary for main-chain folding, stabilization, and dimerization of the domain reside.

Therefore, as was the case for V regions, the least conserved segments of antibody C domains coincide with the loops that interconnect the different β strands in the fully-folded protein. In the case of C region domains, this refers to divergent loops found at each end of the Ig domain, unlike V regions, where the CDRs are clustered at one end of the domain. Not surprisingly, these loops are also where the preponderance of functional interactions have been localized in binding studies. For instance, the binding site on IgG for the FcγRI receptor is located near the hinge region in just such a loop of the $C_\gamma2$ domain. Likewise, the previously mentioned binding of SPA with IgG Fc involves similar loops in both the $C_\gamma2$ and $C_\gamma3$ regions (12). In another example, although no crystal structure is available for IgA, by modeling the IgA sequence on the IgG Fc three-dimensional structure, an exposed loop in the $C_\alpha3$ domain is predicted to be the binding site for the polymeric immunoglobulin receptor.

Thus, the areas of greatest divergence between different C domains are also those implicated in mediating the different biological effects that distinguish one class or subclass of antibody from another. It is fair to speculate that, like the pressures driving the evolution of diversity-generating mechanisms in V regions, similar forces have used the template of the C region Ig domain to select for a variety of distinct binding capabilities and functional attributes, and have utilized a similar region of the domain for these purposes. In a manner analogous to that seen for variable region tertiary structure, those parts of the C region domain that are most malleable structurally are the very same selected evolutionarily for the acquiescence of new biological characteristics.

Quaternary Structure—The Immunoglobulin Monomer

Although, as stated previously, immunoglobulin domains are each capable of folding independently into stable tertiary globular structures, neither individual Ig domains alone nor entire heavy or light chains are ever encountered under normal physiological circumstances. Rather, the simplest form of this protein that occurs naturally is that of the immunoglobulin monomer schematically depicted in Fig. 1. The complete "monomeric" antibody molecule is actually a four-chain dimer of a heterodimer covalently linked by multiple interchain disulfide bonds. In almost all cases, the heavy and light chains are joined by a single cystine to form a "half-monomer" with one complete antigen-binding site, and one or more disulfide bonds in the hinge regions (or hinge domains) link the two heavy chains to form the bivalent tetramer.

Figure 16 represents in two dimensions a member of each of the five classes of human immunoglobulin. Notice that the primary differences between these structures are the presence of the extra C domain in the IgM and IgE isotypes and the number and placement of disulfide linkages and carbohydrate derivatives among the different molecules. Like the intrachain disulfide bond considered in tertiary structure, the interchain cystine attaching heavy and light chains is highly conserved. The cysteines participating in this bond are located at the N-terminal end of C_H1 and the C-terminal end of C_L. IgG1 is an exception where the cysteine donated by the heavy chain is found at the carboxyl end of C_H1 (55). Another exception to the typical bonding pattern is found in the A2m(1) allotype of IgA2. Distinctively, this particular isotype utilizes H–H and L–L disulfide pairing to stabilize the four chains instead of the usual H–L linkage (67). The interchain cystines joining the heavy chains are more variable in both number and position between different isotypes. Generally, H–H bonds are formed between the hinge regions or, in the cases of IgM and IgE, analogous positions in $C_\mu2$ or $C_\epsilon2$ domains, respectively. In addition, both IgM and IgA have two additional half-cystines that have dual roles, depending upon whether the antibody is incorporated into a polymeric form of immunoglobulin. An extra cysteine near the C-terminus is involved in a disulfide bond occurring during J chain–mediated polymerization; another half-cystine in the $C_\mu3$ or $C_\alpha2$ domain forms an intersubunit cystine in the case of IgM or IgA multimers. The disulfide bonds formed by these cysteines in the monomeric forms of IgM and IgA are either interchain (μ and α) or intrachain (α only).

Once interchain disulfide bonds have cemented the four chains into a complete immunoglobulin monomer, the quaternary structure of these molecules can once again be best understood using a domain-by-domain analysis to examine the entire protein. Recall that at this structural level, individual domains interact in such a way that an antibody actually consists of a series of dimeric modules (reviewed in ref. 27). These dimerized domains define the smallest structural and functional units of native immunoglobulin as demonstrated by proteolysis and x-ray diffraction studies. Accordingly, these modules will be reviewed in this context, beginning with the most amino-terminal domain pair, the Fv fragment—or V_H/V_L dimer.

The Fv (see Colorplate 4) is the variable part of the Fab fragment, and as such, constitutes the minimal antigen-binding unit of an antibody. Likely due to this specific functional necessity, V domains dimerize in a manner unlike the strategy employed by all other immunoglobulin domains. In β-pleated sheets, consecutive amino acid side chains protrude on alternating sides at right angles to the plane of the sheet. In most proteins, one side of the sheet packs against another part of the protein (i.e., it is hydrophobic) and the other face of the sheet is exposed to solvent (hydrophilic), leading to a sequence of alternating hydrophobic–hydrophilic residues. In immunoglobulin domains that dimerize, however, this alternating pattern is broken by one of the domain's two β sheets. The sheet that makes up the dimerizing face must interact with both the other β sheet in its own domain and the dimerization face of the adjacent domain. Thus, hydrophilic residues are replaced by hydrophobic side chains that support the dimerization event. While the feature of breaking the alternating hydrophilicity pattern is common to both types of Ig domains that form dimers, V domains interact in a fashion not only specific to V_H and V_L, but also—at

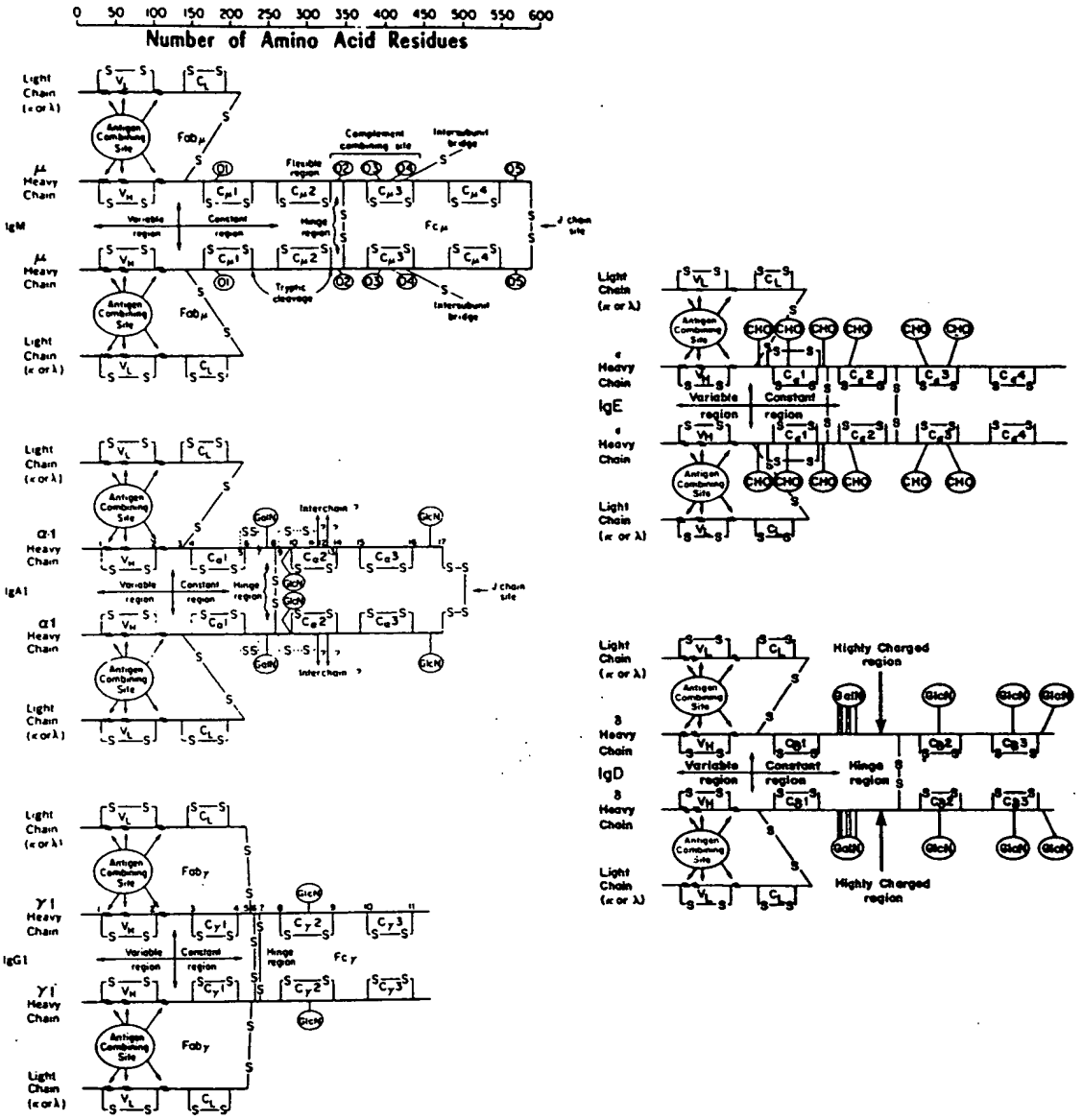


FIG. 16. Schematic representation of the five human immunoglobulin classes. Positions of disulfide bonds and glycosylation are shown for each antibody. (From ref. 66a, with permission.)

the time it was first recognized—unique among all known protein structures (28). Unlike C domains, V domains utilize the five-stranded β-pleated sheet (recall these are unique to V domains) as a dimerization surface (68). The actual V_H/V_L interface consists of the four strands C′-C-F-G on each V domain.

This singular dimerization tactic has profound structural repercussions. In the section detailing tertiary structure, three different β bulges were cited: one common to all V regions found in strand G, one specific to V_H domains located in strand C′, and a third unique to V_L also in strand C′. Note that all of these bulges occur in strands at the edges of their respective β sheets. Normally, β sheets pack such that residues in the middle strands form most of the contacts between layers. On account of these conserved β bulges, however, V domain edge residue side chains protrude into the dimer's interior, preventing a close association of V_H and V_L. The loose packing of the Fv has the effect of creating a hydrophilic groove into which small molecules can fit. This groove, which is primarily lined by residues of the HVRs, and the remaining CDR loops at the end of the V regions form a potential antigen-binding site whose surface area is approximately 2000 to 3000 Å^2 (69).

Many of the residues consistently responsible for V_H/V_L interdomain contacts have been localized. About half of the hydrophobic core contacts are formed between FR2 of one chain and FR4 of

the other. The majority of remaining interactions involve CDR3 of one chain and the FR2 and/or CDR3 of the complementary domain. Overall 12 to 21 V_L and 16 to 22 V_H residues participate in interchain stabilization (69). Given the extensive contribution of HVRs to these interactions (28), one might expect that the affinity of H–L pairing might also be variable. Nonetheless, a number of studies using heavy and light chains from different individuals, or even different species, have demonstrated the capacity of heterologous H–L pairs to form stable associations. This implies that the conservation of the basic structural features of V_H and V_L domains has persisted throughout evolution, or at the very least not diverged to the point where the C_H1 and C_L elements are unable to anchor productive interactions between these disparate entities.

Between the Fv and Fb (C_H1/C_L dimer) fragments, a short polypeptide stretch exists that is vital to both the Fv's ability to bind antigen productively and to the ability of C_H1/C_L domains—and all constant domains C-terminal to them—to dimerize properly. This region, comprising the carboxy-terminal amino acids of the V region contiguous with the N-terminal residues of the C region, connects Fv to Fb in the complete Fab fragment and is known as the *elbow peptide.* Collectively, the two elbows of an immunoglobulin Fab are also referred to as the *switch peptide.* Several Fab crystal structures have demonstrated the switch to be a flexible segment permitting considerable bending between the V and C domains (70). This is thought to be important in allowing Fabs to bind epitopes of varying spatial arrangement. An equally important feature of the individual elbow peptides is the fact that they make possible a remarkable 180-degree rotation in the quaternary structure of antibody domains that is essential for the correct orientation of all C domains in order for dimer formation between them (8).

These 180-degree rotations occur at the elbows between V_H-C_H1 and V_L-C_L, and are necessary to properly position the C regions. The most N-terminal C domains, C_H1 and C_L, are then able to combine to form the Fb fragment. C_H1 and C_L are prototypes for the C type Ig domain. Like all C domains, they lack the C′ and C" strands present in the V domain five-stranded sheet. Like V domains, they also break the alternating hydrophilicity pattern in one of their two β sheets, but in the case of these domains, this occurs on the four-stranded face of the immunoglobulin fold (A-B-E-D) instead. As a result—and due to the permissive rotations at the elbows—C_H1 and C_L utilize the opposite (relative to V_H and V_L) sides of their domains to dimerize (68). The less-polar residues at the core of these dimeric C modules are conserved in both C_H1 and C_L across species (8–10,12), and in this case (unlike V domains), they tend to reside primarily in the middle strands (B and E), where they mediate a much "tighter" association. Consequently, the Fb is often perceived as a compact anchor for the V domains, forming a stable platform upon which antigen binding can occur. Fb, together with the hinge, also serves as a spacer between the Ag-combining site and the bulky Fc region.

Between the Fab and Fc regions, immunoglobulin hinges (or extra domains) are critical determinants of overall antibody structural and functional properties. Structurally, hinges are extended segments of dimeric peptide held together by one or more disulfide bonds and dominated by prolines, serines, and threonines. This amino acid composition gives hinges flexibility and gives the Fab arms of an antibody flexional (71) and torsional (72) mobility. Hinge flexibility allows the Fabs to conform to the arrangement of epitopes in order to bind bivalently—presumably giving an antibody greater avidity and versatility. The degree of flexing permitted correlates strongly with hinge length between the end of C_H1 and the first interheavy disulfide bridge (the "upper hinge") (71,73); thus IgG3 is more flexible than IgG1. Hinge length and flexibility also reduce the steric barrier that Fabs may present to the access of C_H2 by other molecules. For example, while normal human IgG1 activates complement effectively, the hinge-deleted variant IgG1 paraprotein *Mcg* is unable to fix complement because the Fabs rest too close to the Clq-binding site on C_H2 to allow interaction.

The hinge, being an extended peptide, is the most proteolytically labile part of an immunoglobulin; recall that the early studies that resulted in the understanding of Fab and Fc relied on proteolytic digestion of the hinge. For instance, the δ isotype, with its long, charged hinge, is very susceptible to proteolysis, which may explain its short serum half-life (74). This issue is of critical importance to IgA, which serves its primary function at mucosal surfaces where proteases from bacterial and host sources are prevalent. For example, the α1 hinge contains five carbohydrate attachment sites within a stretch of only 17 amino acids (75), rendering IgA1 resistant to cleavage by intestinal proteolytic enzymes. However, several strains of bacteria secrete proteases that specifically target the α1 hinge. Presumably as an evolutionary consequence, the hinge of IgA2 has undergone a 13-amino acid deletion, restoring its resistance to this second form of proteolytic challenge (53,76,77). Similarly, structural features unique to hinges of each of the different isotypes may have evolved so as to confer their respective immunoglobulins with specific functional characteristics.

C-terminal to the hinge, the Fc region resides. In the cases of IgG, IgA, and IgD, the Fc is a dimer of two C_H2–C_H3; in IgM and IgE it consists of paired C_H2–C_H3–C_H4 domains. Structurally, $C_{\mu/\varepsilon}3$ is equivalent to $C_\gamma2$, and $C_{\mu/\varepsilon}4$ is homologous to $C_\gamma3$. As has been described above, the vast majority of sites that define the function and physiology of a particular isotype map to the Fc region. The first striking quaternary structural feature of the Fc is that $C_\gamma2$ (and its structural homologues) fail to dimerize. Analysis of these regions demonstrates that C_H2 domains possess a hybrid structure intermediate between V and C domains. Whereas V domains are five-strand/four-strand sandwiches (dimerizing on the five-strand face) and C domains are three-strand/four-strand sandwiches (dimerizing on the four-strand face), $C_\gamma2$ is a four-strand/four-strand sandwich. Moreover, substitutions present in outward-pointing side chains on both sides of the domain prevent dimerization along either face (11). Another mixed feature observed is that the β strands of $C_\gamma2$ are of lengths longer than those of V domains, yet shorter than those of either $C_\gamma1$ or $C_\gamma3$ domains. Finally, all C_H2 domains are derivitized by an N-linked oligosaccharide near the middle of the domain (refer back to Fig. 1), except $C_\alpha2$, where it is found nearer to the carboxy-terminal end of the domain. Hydrogen bonding between these sugars serves as the only contact between C_H2 domains. Moving C-terminally, another important distinction concerning C_H2 is that longitudinal contact with the C_H3 domain (about 340 Å^2 surface area/domain) prevents little interdomain bending (12), unlike the flexible elbows between Fv/Fb and the hinge separating Fab and Fc.

The final domains of the Fc, $C_\gamma3$ and its structural equivalents, pair in the manner described for C_H1 and C_L. Studies using limited proteolysis, reduction, and denaturation originally designated this fragment, a dimer of C_H3 domains, as pFc′. C_H3 domains, like the Fb, dimerize with tight association between them (1100 Å^2/domain), using the four-stranded faces of each domain (11). Also like C_H1 and C_L, all C_H3-domain isotypes show conservation of core residues involved in this pairwise domain interaction.

Structurally, the only feature that makes a marked distinction between the different C_H3 domain homologues is the presence of the 18-amino acid "tail-piece" that exists at the carboxy-termini of the $C_\alpha3$ and $C_\mu4$ domains. This sequence is important to polymerization and will be discussed further in the section on higher-order immunoglobulin structure. Taken together, the Fc can thus be thought of as an approximation of the Fab, having two pseudo-V regions (the C_H2s) at its amino-terminus, and a module of C-terminal constant regions (the C_H3s) dimerized in the mode characteristic of classic C-type domains.

Although similar in many ways at the protein level, one of the properties that most notably differentiates the quaternary structures of the five classes of Fc is their pattern of glycosylation—with significant functional ramifications. For instance, while the oligosaccharide moiety of the IgG molecule accounts for only 2% to 3% of its mass, it has been shown to be essential for optimal activation of effector mechanisms leading to the clearance and destruction of pathogens (78–80). As introduced above, all human antibody molecules of the IgG class have N-linked oligosaccharide attached at the amide side chain of Asn 297 on the β-4 bend (between β strands D and E) of the inner face of the C_H2 domain of the Fc region (81). This oligosaccharide moiety is of the complex biantennary type, having a hexasaccharide "core" structure (GlcNAc2Man3GlcNAc) and variable outer arm "noncore" sugar residues, such as fucose, bisecting N-acetylglucosamine, galactose, and sialic acid (see Fig. 17). In all, a total of 36 structurally unique oligosaccharide chains may be attached at each Asn 297 residue. It is likely, but not certain, that the precise fidelity of this glycosylation is important.

The site for this C_H2 carbohydrate is a conserved feature for all mammalian IgGs, and glycosylation occurs at a homologous position in human IgM, IgD, and IgE molecules. As stated above, IgA is also glycosylated within its $C_\alpha2$ domain, but at a site further C-terminal. Human IgM, IgA, IgE, and IgD molecules also bear additional N-linked oligosaccharide moieties attached to the C domains of their heavy chains. Furthermore, IgA1 and IgD proteins also possess multiple O-linked sugars in their extended hinge regions, attached to the hydroxyl groups of serine and threonine residues. Glycosylation, in one form or another, is in fact characteristic of all heavy chain C regions and remains one of the most active areas of research in immunoglobulin structural biology.

The structural and functional consequences of Fc oligosaccharides have begun to be assessed experimentally by comparison of glycosylated and aglycosylated forms of IgG. The latter is ordinarily generated by growing IgG-producing *E. coli* in the presence of tunicamycin (a glycosylation inhibitor) or by protein engineering of the carbohydrate acceptor sequence. One characteristic apparently affected by the sugar moieties found on antibodies is the duration of these proteins' existence *in vivo*. Studies of the blood clearance of glycosylated and aglycosylated mouse/human chimeric IgG1 molecules in mice reveal accelerated clearance for the aglycosylated form despite similar half-lives. Additionally, galactosylated and agalactosylated IgG have been investigated to determine the role of outer-arm sugars in complement (C1q)-mediated lysis. The agalactosylated form, produced following exposure to β-galactosidase, has an observed two fold lower activity than the galactosylated form (82). Related studies have substantiated these conclusions (83).

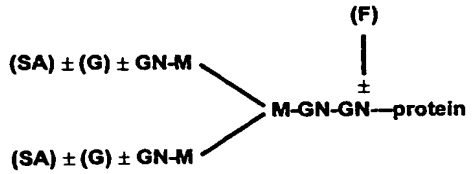


FIG. 17. Schematic representation of N-linked sugars attached to all C_H2 domains. The core carbohydrate moiety of the complex form of oligandosaccharides is represented by the sugar residues in *open type*. The possible outer-arm residues are *in parenthesis*. All possible combinations are observed. SA, sialic acid; G, galactose; GN, N-acetylglucosamine; M, mannose; F, fucose. Attachment of oliosaccharide occurs on the amide side chain of the Asn-X-Ser/Thr sequon (X ≠ Pro). The Ser/Thr residue forms hydrogen bond(s) with the amide group in order to activate it for attachment to the primary *N*-acetylglucosamine residue of the dolichol intermediate (catalyzed by oligosaccharyltransferase). (From ref. 81a, with permission.)

The vital importance of correct glycosylation was further provided by a study using a human/mouse chimeric IgG1 molecule produced in yeast cells, and anticipated to have high mannose-type oligosaccharide attached at Asn 297. The yeast IgG1 product was unable to activate C1q to trigger human complement-mediated lysis of targets, while the same chimeric IgG1 construct expressed in rodent cells (and therefore glycosylated normally) was effective in that regard. A direct role for oligosaccharide in activating the complement cascade is apparent with mannan-binding protein, a lectin that can function as a surrogate C1 component. The specificity of mannan-binding protein is for mannose and N-acetylglucosamine residues, and it has been shown that it can access and bind to terminal N-acetylglucosamine residues exposed on agalactosyl IgG (84).

Studies utilizing the three types of human Fcγ receptors (FcγRI, FcγRII, and FcγRIII) have also attested to the significance of oligosaccharide modifications on antibodies. The IgG subclass specificity of the FcγR suggests that recognition is correlated with the presence or absence of carbohydrate derivatives. This conclusion is supported by the demonstration that aglycosylated human chimeric IgG3 has a reduced interaction with all three Fcγ receptors. Moreover, at the level of function, while haptenated RBCs sensitized with this same aglycosylated IgG3 antibody could still trigger superoxide production by U937 cells, higher levels of sensitization were required compared to normally glycosylated IgG3. The aglycosylated IgG3 also was not recognized by human FcγRII expressed on K562 and Daudi cells, had reduced rosette formation (mediated through FcγRII expressed on human NK cells), and essentially abolished antibody-dependent cellular cytotoxicity (85,86). Clearly these conclusions illustrate that proper Fc glycosylation is—at least in some cases—necessary for normal structural recognition and biologic function of immunoglobulins.

Glycosylation is potentially important outside of the Fc region as well. It has been estimated that up to 30% of polyclonal IgG molecules are also derivitized by oligosaccharide within the Fab region. Since there are no known sites for sugar attachment in $C_\gamma1$ or C_L, this is most likely in the V regions. Of interest in this regard, an analysis of the DNA sequences of 83 functional human germline V_H gene segments revealed five that encoded potential glycosylation sites. Some, but not all, of these are known to be glycolysated. In one study of protein and cDNA V_H and V_L sequences, about 25% had potential glycosylation sequences, some of which had arisen as a result of somatic mutation and antigenic selection (87). In most circumstances of V region glycosylation studied thus far, the oligosaccharide moiety does not contribute directly to ligand binding, but can exert a subtle influence on protein tertiary and quaternary structure that is essential for full activity of the antibody. Thus,

oligosaccharides occur in many places on immunoglobulin molecules and can affect antibody characteristics as disparate as antigen-binding and the assortment of different Fc-associated functions.

Higher-Order Immunoglobulin Structure—Polymeric Immunoglobulin

One of the most fascinating structural attributes of immunoglobulin is the ability of two classes of heavy chain, IgM and IgA, to form higher order multimeric complexes. IgM and IgA antibodies do not always form polymers, however; monomeric α and μ isotypes exist in forms analogous to those for the γ, δ, and ε isotypes, as well. In addition, polymeric immunoglobulin (pIg) can typically come in a variety of manifestations. The most common forms of these molecules are dimeric (IgA) and pentameric (IgM), although other polymers have also been described. Electron micrographs of murine pentameric and hexameric IgM and human dimeric and trimeric IgA molecules are presented in Fig. 18. Multimerization obviously increases the number of potential antigen-binding sites available on the antibody, and this increase in valence translates into enhanced avidity for polymeric, low-affinity epitopes. This is particularly beneficial for antibodies of the μ and α classes, which serve as the first line of defense at mucosal surfaces where encounter with this type of pathogenic target (i.e., cell-wall polysaccharides) is most frequent. Moreover, IgM, which is the antibody characteristic of primary humoral responses when affinity maturation has not yet occurred, is reliant upon this increased avidity to mediate its functional responsibilities. In addition to raising the apparent affinity for antigen binding, polymeric IgM's juxtaposition of several Fc regions in close proximity also likely plays a role in its efficacy in fixing complement components of the classical pathway. Mechanistically, assembly and secretion of pIg involves the covalent linkage of concatomers of prototypic immunoglobulin monomers, and two accessory proteins termed *J* (joining) *chain* and *secretory component* (SC) play key roles in these processes.

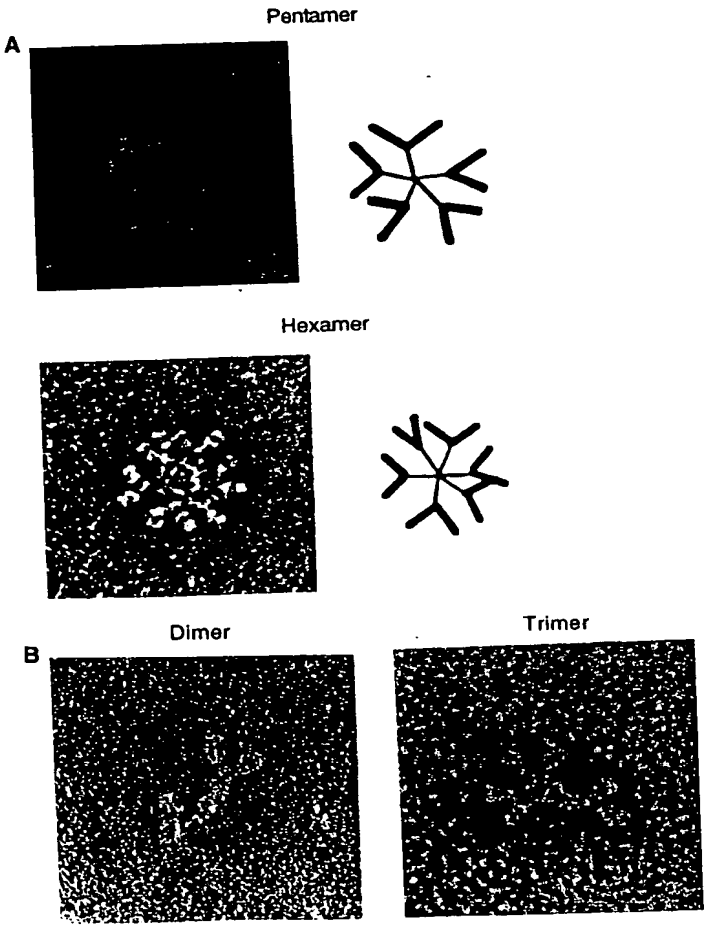


FIG. 18. Electron micrographs of immunoglobulin multimers. In (**A**), a murine IgM pentamer and interpretive diagram (*top*) and a murine IgM hexamer and diagram (*bottom*) are displayed. In (**B**), a human IgA dimer (*top*) and trimer (*bottom*) are presented. All magnifications are ×600,000. (From ref. 87a, with permission, and courtesy of K. H. Roux.)

J chain is a 137-amino acid polypeptide synthesized by pIg-producing plasma cells. J chain covalently interacts with one or more cysteines of immunoglobulin monomers undergoing multimerization (88). It is a proteolytically labile molecule with a high content of negatively charged amino acids and has eight cysteine residues involved in both intra- and interchain disulfide bonds (89). The high level of conservation between J chains of human (90), murine (91), rabbit (92), and even amphibian (93) origin implies that there is a powerful selective pressure to maintain J chain structure. A report identifying J chains in invertebrates which have no known correlate to antibody proteins (94) also indicates that the J chain probably performs some basic protein function that predates its eventual development of the ability to interact with immunoglobulin. In any event, structural studies imply that despite the fact that J chain lacks any significant sequence homology with immunoglobulin, it likely folds into a β-barrel structure similar to that of an immunoglobulin fold (95). Besides the intrachain cystines which stabilize the J chain itself, additional cysteine residues form disulfide bridges to the tailpiece of one or more immunoglobulin monomers during multimer assembly (89). Although the actual details of polymerization have not as yet been elucidated, it is known that the 18-amino acid $C_{\mu/\alpha}$ tailpiece and its penultimate cysteine residue are necessary for the process (52,96).

The stoichiometry of multimer assembly is such that one J chain is present per polymer (whether dimer, trimer, pentamer, etc.). While not always a part of pIgM molecules, J chain is probably absolutely necessary for formation of polymeric IgA, as it is always present in the complex. J chain synthesis is known to be highly regulated (97), and it is thought to be linked to the B cell's activation state as well (98). High levels of J chain expression have been shown to result in production of normal J chain–associated pentameric IgM, while low J chain synthesis results in secretion of hexameric IgM lacking the protein. Intriguingly, this hexameric IgM is actually 20-fold more potent at promoting lysis by complement than is the usual pentamer (99).

The second accessory molecule associated with the secretion of multimeric antibodies actually derives from another protein belonging to the IgSF—the pIg receptor (pIgR)—and is not even made by cells of the B lineage. Secretory component (also called *secretory piece*) was initially discovered as a polypeptide found tightly complexed to the Fc of secreted forms of IgA and IgM (100); subsequently, it was recognized that SC is actually a portion of the larger transmembrane pIgR protein (101). The entire cDNA sequence of the pIgR reveals a protein consisting of seven domains: the first five are extracellular and structurally similar to

immunoglobulin V regions, the sixth contains a transmembrane segment and is partially homologous to immunoglobulin V domains, and the seventh contains an unrelated C-terminal intracellular domain (102). The first five domains of the pIgR are in fact the secretory piece originally co-isolated as part of the secreted immunoglobulin complex.

The pIgR is synthesized in epithelial cells of the respiratory, gastrointestinal, and genitourinary tracts and is expressed on their basolateral aspect, where it binds to pIgA and pIgM in a high-affinity interaction. It is known that the N-terminal domain of the pIgR confers binding specificity, and it is thought that both J chain and Fc $C_\alpha 3/C_\mu 4$ determinants are recognized by the receptor. Interestingly, although the precise molecular locations of pIg/SC interaction have not been identified, there is evidence that, at least structurally, the site is well conserved. Studies have shown, for instance, that human SC binds not only human pIgA and pIgM (103,104), but also several other mammalian species' IgM and IgA (105), and even chicken IgA (106). This cross-reactivity, however, may be mediated by the J chain rather than the Fc regions.

Regardless, following the initial interaction between domain 1 of the pIgR and the C-terminal domain/J chain of the pIg, secondary contact occurs between pIgR domains 3, 4, and 5 with the antibody, consumated by formation of a disulfide linkage between the SC and $C_\alpha 2/C_\mu 3$. This covalent bond is between Cys 467 in domain 5 of the pIgR and Cys 311 located in the $C_\alpha 2$ domain of one IgA subunit's heavy chain (an IgA dimer would have four $C_\alpha 2$ domains overall) (107). After a stable interaction has been established, endocytosis of the complex occurs via clathrin-coated pits. Next, following cleavage between domains 5 and 6 of the pIgR, the poly-Ig /SC (now formally termed *secretory immunoglobulin*) is exocytosed at the apex of the cell, releasing the secretory immunoglobulin onto the mucosal surface (108). Current thinking holds that SC's most important function, outside the realm of its role as the pIg-binding portion of the pIgR, is to help protect secretory immunoglobulin in harsh mucosal environments.

Structurally then, polymeric antibodies represent the pinnacle of complexity in terms of immunoglobulin's expansion upon the Ig homology domain concept. From the fundamentals of a simple 110-amino acid domain with a few conserved core residues and a basic structural topology, an intricate molecule such as pentameric IgM (containing 70 different Ig domains of both V and C types, not to mention Ig domains contributed by SC!) is constructed. A molecule capable of recognizing as many as ten (although steric constraints usually dictate less) identical specific antigenic determinants, and also able to mediate several different important biologic functions—all of which will be detailed in the following section.

IMMUNOGLOBULIN FUNCTION

Throughout this chapter thus far, many of the differing functional capacities of antibodies have already been alluded to, as pertains to the identification of the specific structural determinants responsible for particular interactions. Still, the plethora of biologic activities performed by immunoglobulin is best treated as a separate section, in which the many and varied aspects of immunoglobulin function can be detailed and integrated in a physiologic context. Collectively, secreted antibodies are able to activate both the classical and alternative complement cascades (see Chapter 29), transcytose across epithelial cell layers to provide a barrier to pathogens at mucosal surfaces (see Chapter 27), travel transplacentally to confer maternal humoral immunity to the fetus and neonate, induce phagocytosis by macrophages and granulocytes via the process of opsonization (see Chapters 30 and 41), foster antibody-dependent cellular cytotoxicity by lymphocytes and NK cells (see Chapters 17 and 31), encourage antiparasite immune responses by eosinophils (see Chapter 38), and promote degranulation by mast cells and basophils (see Chapters 32 and 35)—not to mention antibody's ability to bind and inactivate foreign antigenic entities directly (see Chapter 39)!

Even this imposing list of attributes neglects to mention the myriad effects mediated by surface immunoglobulin that include, but are not limited to, the induction(s) of activation (see Chapter 7), differentiation (see Chapter 6), anergy (see Chapter 20), and even apotosis (see Chapter 23) of B lymphocytes, which are detailed elsewhere in this volume. Surface Ig on memory B cells also has the ability to act as a high-affinity receptor for the recognition, internalization, degradation, and eventual presentation of specific antigens to T cells (see Chapter 9). This allows memory B cells to act as *antigen-specific* antigen-presenting cells (APC), which makes them uniquely efficient among this class of cells. Moreover, emerging fields of study, such as the growing body of literature concerning intracellular antibodies, indicate that new functional capacities for immunoglobulin are likely yet to be discovered.

As for the preceding sections on immunoglobulin structure, the biologic capabilities of immunoglobulin are best treated by dissecting the molecule into the portions responsible for each of its different functional characteristics. While there are some exceptions, in general, specific functions of antibodies can be ascribed to individual domains of the molecule. In the case of variable regions, this requires consideration of the two V domains (V_H and V_L) whose primary function is the binding of antigen. Additionally, it has also been appreciated that certain "superantigens" bind to the V domains as well. In the case of constant regions, because no effector properties have been linked to C_L domains, this entails discussion of each of the heavy chain isotypes (IgM, IgD, IgG, IgA, and IgE), whose functional differences must be a direct result of their structural heterogeneity.

Variable Region Functions

The two V regions (either V_H/V_λ or V_H/V_κ) together form the variable domains of an antibody molecule and provide the specificity for targeting the effector arms of immune response. In general, both V regions are needed to provide specificity and high affinity. There are many examples of individual variable regions binding antigen, but clearly, when the two chains act in concert, the exquisite specificity and affinity of interaction between antibody and foreign antogen is dramatically enhanced.

The concept of hypervariable regions and complementarity-determining regions is pertinent here. In general, most of the contacts between the V domains and antigen take place between amino acid residues in the CDRs and the major epitopes on the antigen (109). However, more recent studies have documented considerable contact between so-called framework residues and antigen. This is most dramatically seen in the lysozyme–antibody crystal (see Colorplate 6) where many residues (especially in the heavy chain FR3 region) are in direct contact with the antigen (110). The generalization that the CDRs provide *all* of the contact residues grew out of the early work involving hapten/anti-hapten systems. When only

small organic haptens are the "antigen," then the CDRs can easily provide a "pocket" in which antigen engages antibody (recall that the peculiar dimerization strategy employed by V domains has the propensity to generate such pockets). Closer inspection of the lysozyme Ag-Ab complex—in particular, lysozyme residue Gln 121—is instructive in this regard. Gln 121 protrudes into the cleft between V_H and V_L much like haptens fit into the groove described above. However, other non-groove residues still appear to provide the bulk of the interacting surface for this antibody. In fact, other antibodies that are also reactive with lysozyme (see Colorplate 6) have further borne out this face-to-face binding concept (59). It is fair to say, then, that generally when large molecules such as proteins complex with antibody, the interaction is one of two protein "faces" coming together. In that instance, the notion of a pocket is less appropriate, and as such, non-CDR residues (especially in FR1 and FR3) are also frequently involved.

As with all molecular associations, antigen–antibody interactions occur only if the binding reaction releases enough free energy to be thermodynamically favored. The affinity of interaction is exponentially related to changes in free energy (see Chapter 4). Free energy changes are the sum of changes in both entropy and enthalpy, with increases in entropy and decreases in enthalpy favoring binding. Few association reactions are able to fulfill both of these requirements, however. Instead, a favorable change in one component compensates for a less unfavorable change in the other. When antibodies bind their ligands, the freedom of one molecule to move relative to the other is lost (an unfavorable decrease in entropy). Stabilization of most conformational motions of both the epitope and the backbone and side chains of the paratope surface lowers entropy even further. Thus, to encourage binding to antigen, antibodies must attempt to limit decreases in entropy and offset these losses by potentiating decreases in enthalpy. At the amino acid level, this leads to a selection for Tyr, Trp, Ser, and Asn in combining sites (111,112), because these residues have lower conformational freedom, and hence less entropy to lose upon binding. Additionally, the side chains of these residues foster the varied chemical interactions that drive changes in enthalpy necessary to promote binding energetically.

Specifically, the antigen–antibody interaction involves a variety of forces, including electrostatic (the attraction between opposite charges), hydrogen bonds (hydrogen shared between electronegative atoms), van der Waals forces (the fluctuations in electron clouds around molecules oppositely polarize neighboring atoms), and hydrophobic forces (hydrophobic groups interact unfavorably with water and tend to pack together to exclude water molecules) (113). Of course, salt bridges and other forms of interaction also come into play in some specific immunoglobulin–ligand complexes as well. It is also important to appreciate that rarely do covalent bonds occur between antigen and antibody. Thus, antigen–antibody complexes are readily dissociated by solvents that break the above bonds, such as high salt, organic solvents, urea, and so on.

Thermodynamic considerations for ligand binding are favored by large interacting surfaces of both antibody and antigen, which are packed as closely as possible. Large interaction areas exclude more bound water, somewhat opposing losses in protein entropies with gains in solvent entropy. Surprisingly then, some antigen–antibody complexes actually retain water molecules in their interfaces. However, rather than interfering with binding, these frequently contribute to the interaction by hydrogen bonding to both surfaces (114). More important than entropic changes, the overriding impetus for large contact areas of antibody and antigen is their ability to bring about large decreases in enthalpy. This effect is mediated by allowing many chemical interactions of all kinds to occur simultaneously between the epitope and paratope.

A model for interaction outside the realm of typical antibody–antigen binding has recently come into vogue, that of the "superantigen." Superantigens were first appreciated in the context of T cell activation. Certain molecules (particularly bacterial products) were found to interact with many different T cell receptors (TCRs) having a variety of specificities (115). Thus, superantigens were originally defined as intact proteins that stimulated large numbers of T cells by binding the V region of a specific family of V_β chains (the heavy chain of the TCR) outside its normal binding groove. Typical T cell superantigens simultaneously stimulate 5% to 25% of the T lymphocyte population, compared with 0.01% stimulation of T cells by a conventional antigen.

More recently, this concept has been extended to B cells (116,117). SPA is a prototype of a B cell superantigen. Although SPA was known to bind to certain immunoglobulin C regions—it has long been used as a mitogen for human B cells—it has been shown that SPA also binds to certain human V_H3-encoded antibodies (118,119). It also binds to the Fab region of murine immunoglobulins, particularly those of the J606 and S107 V_H gene families (which belong to the same clan as human V_H3). SPA binds independently of D, J_H, and light chain utilization, although some light chains influence the extent of binding. Since the interaction is independent of the specificity of the antibody, and since SPA in general does not block antigen binding, it is considered a B cell superantigen. SPA is even able to deliver activation signals to stimulate the differentiation of those B cells containing V_H3-encoded receptors, and SPA also stimulates antibody production. More recent work (120) has documented that SPA simultaneously interacts with FR1, CDR2, and FR3 on the V_H region (Fig. 19).

Other superantigens have also been described that are able to bind immunoglobulin in regions apart from the traditional antigen-

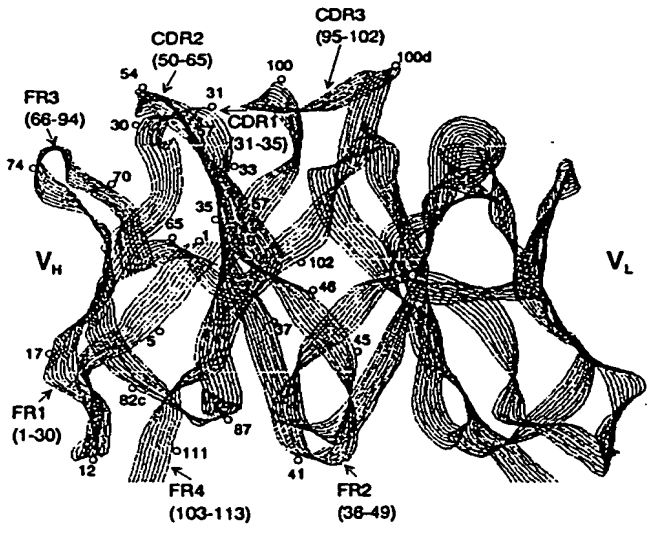


FIG. 19. Ribbon drawing of the Fv fragment of the V_H3 antibody KOL. The FR1, CDR2, and FR3 subdomains of the heavy chain (*left*) are juxtaposed in a manner forming a solvent-exposed face which allows SPA binding. (From ref. 120, with permission.)

combining site (based on their broad specificities). These include the HIV envelope protein gp120 (121), which like SPA binds V_H3-encoded antibodies, and the TCR-associated molecule CD4 (122). Like the circumstance with antigen binding, superantigens generally require both V_H and V_L domains (even though the particular identity of the light chain is unimportant). Individual heavy chains do not bind the B cell superantigens that have been described to date, indicating that light chains must influence their conformations appreciably. Finally, the report of a crystal structure (an IgM rheumatoid factor Fab complexed to its autoantigen, an IgG Fc) showing residues at the edge of the conventional binding site mediating interaction indicates that still more novel paradigms for antibody–antigen binding possibly exist as well (123).

Constant Region Functions

Because mammalian species each utilize the same major classes of antibody (although their organization of subclasses differs), it is reasonable to presume that each isotype subserves some vital biologic function(s). Along these same lines, it should be remarked that even in "lower" species, where only one type or one copy of heavy chain gene is present, the protein product resulting from this element is typically heterogeneous. In other words, although fish, amphibians, and reptiles all possess fewer immunoglobulin isotypes than do mammals at the genomic level, greater than one C region protein is produced per gene. For example, sharks make both monomeric and polymeric IgM; in skate, turtle, and duck there are truncated and full-length versions of the immunoglobulin polypeptide; *Xenopus* immunoglobulin comes in both glycosylated and aglycosylated forms (51). Clearly, evolution has recurrently employed the strategy of adopting more than one type of antibody to perform the multitude of biologic responsibilities that are required by species for effective immunologic functioning.

In a broad sense, Fc-mediated effector functions can be classified into three general categories: (a) activation of complement, (b) interaction with effector cells, and (c) transport and compartmentalization of immunoglobulins. In addition, different isotypes have different stabilities *in vivo*, such that this is an important variable as well. In the following sections, the five classes of human immunoglobulins are each discussed separately with respect to function. Table 2 presents a summary of key properties for each class of human antibody, and Fig. 20 compares the circulating serum levels for each of the five major isotypes.

IgM

IgM is the most versatile antibody and almost certainly the first type of immunoglobulin to have developed evolutionarily. Heavy chains of the μ class are the first type expressed during B cell development, and IgM is the isotype produced in primary immune responses. IgM, in the form of surface immunoglobulin, is also an important receptor on immature B lymphocytes and on mature, naive B lymphocytes. Total serum Ig consists of 5% to 10% IgM, and second to IgA it is the major isotype of mucosal immunity. Originally named due to their description as *m*acroglobulins, IgM molecules are thought to serve similar functions in all mammalian species. In fact, IgM-like (polymeric, having five domain heavy chains with large carbohydrate content, and present as a cell surface receptor on most B cells) antibodies have even been identified in most non-mammalian vertebrates other than the jawless fish (51). Unquestionably, the polymeric structure of IgM has been conserved in evolution, probably due to its higher avidity for antigen compared with that of the monomer.

TABLE 2. *Physical, chemical, and biological properties of human heavy chain immunoglobulin classes*

Property	IgM	IgD	IgG	IgA	IgE
Molecular form	Pentamer, hexamer	Monomer	Monomer	Monomer, dimer	Monomer
Number of C region domains	4	3	3	3	4
Tailpiece	+	–	–	+	–
Accessory chains	J chain, SC	None	None	J chain, SC	None
Subclasses	None	None	G1,G2,G3,G4	A1,A2	None
Molecular weight	950 kD, 1150 kD	175 kD	150 kD	160 kD, 400 kD	190 kD
Carbohydrate content (%)	10	9	3	7	13
Percentage of total serum Ig	5–10%	0.3%	75–85%	7–15%	0.02%
Average adult free serum level (mg/ml)	0.7–1.7	0.04	9.5–12.5	1.5–2.6	0.0003
Synthesis rate (mg/kg/d)	7	0.4	33	65	0.016
Serum half-life (d)	5	3	23	6	2.5
Antibody valence	10, 12	2	2	2, 4	2
Bacterial lysis	+	?	+	+++	?
Placental transfer	–	–	+	–	–
Mast cell/basophil binding	–	–	–	–	+
Macrophage binding	–	–	+	+	–
Classical complement activation	++	–	+	–	–
Alternate complement activation	–	+	+	A1+,A2–	–
Other biological properties	Primary Ab responses; Secretory immunoglobulin	Unknown; Useful as a B cell marker	Hallmark of secondary immune responses	Main secretory immunoglobulin	Allergic and anti-parasite responses

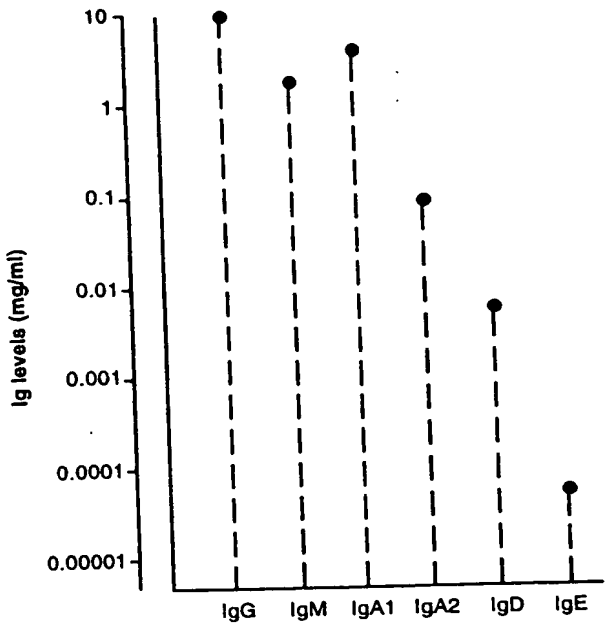


FIG. 20. Circulating levels of different human immunoglobulin isotypes. Note the log scale of the graph and that both human IgA isotypes are represented. (From ref. 123a, with permission.).

The two most common forms of IgM are the membrane-bound monomeric form and the secreted pentamer. The cell-surface version of IgM serves as the antigen-specific receptor for B cell activation, although the activation signal is actually transmitted by the transmembrane accessory molecules Igα and Igβ (124). It is unclear whether the $C_\mu 4$ domains of surface IgM participate in the interaction with the α/β heterodimer, but it is more likely that the important associations lie in the transmembrane and cytoplasmic regions of IgM that are specific to the cell-surface form (125). The surface form of IgM is also important in the development of the B cell. During pre-B stages, μ heavy chains are associated (via a disulfide linkage in $C_\mu 1$) with the "surrogate" light chains $V_{pre}B$ (analogous to a V_L domain) and λ5 (a C_λ analogue) (126,127). This complex, once again through accessory molecules, is able to transduce signals thought to be necessary for allelic exclusion of the other heavy chain locus and for subsequent light chain rearrangment (detailed in Chapter 5). Eventually, of course, the IgM heavy chains become associated with either λ or κ light chains.

Polymeric IgM also has its own catalogue of functional attributes. IgM antibodies are the first to be secreted from plasma cells upon challenge by antigen; since IgM is not secreted in large quantities from memory B cells, elevated IgM is indicative of recent antigenic exposure. As stated earlier, IgM antibodies generally have low affinity, as they have not gone through the processes of somatic hypermutation and affinity selection. Nonetheless, the high avidity of pIgM renders it capable of efficiently binding antigen. Similarly, a single polymeric IgM molecule is able to effectively initiate classical complement fixation, even though the affinity of C1q for Cμ is very low. The C1q-binding site of IgM has been localized to the $C_\mu 3$ domain (128) and appears to be dependent upon carbohydrate found there for potent binding (129). While $C_\mu 3$ domains (and their structural homologues) are not well conserved evolutionarily, the ability to mix IgM and complement from different species and retain activity indicates that the complement recognition sites on vertebrate immunoglobulins may be similar. In fact, even *Xenopus* IgM has been demonstrated to fix mammalian complement components (130)! IgM has also been shown to interact with C3b via its $C_\mu 1$ domain, thereby allowing antibody–antigen complexes containing IgM to indirectly mediate phagocytosis. By this mechanism, C3b, once fixed, can promote uptake via complement receptors found on macrophages.

The high avidity of IgM for both antigen and complement is crucial in the context of its role as a front-line defense mechanism. IgM not only is the humoral agent of primary immune responses, but also—like IgA—is transported by the pIgR across epithelia such that it serves a role as a secretory immunoglobulin at mucosal surfaces. Since secretory immunoglobulins are present in breast milk as well, IgM also participates significantly in protecting the newborn from intestinal pathogens until such time as the neonatal immune system is fully functioning. A role for IgM in mucosal immunity must have developed early in evolution, as it is the sole immunoglobulin in some animals.

IgD

IgD is present in serum in very low amounts (less than 0.5% of total serum Ig). Although synthesis of IgD is also very low (at least an order of magnitude lower than that of IgM, IgG, and IgA), IgD's pronounced susceptibility to proteolytic degradation is probably also responsible for its scarcity in plasma and other bodily fluids. The unusually long hinge region linking Fab to Fc in IgD is thought to be largely accountable for its short half-life. IgD is secreted neither during an immune response, nor following mitogenic stimulation of IgD+ B cells, although in the form of immune complexes it is known to be able to activate the alternative complement cascade. IgD's low levels make this complement fixation unlikely to be important in the *in vivo* context. In fact, no specific functions unique to IgD have been definitively assigned to the δ Fc region in either its membrane-bound or soluble forms. That notwithstanding, the fact that the IgD class is maintained in all mammals, has a high level of conservation across species (131), and the existence of an Fcδ receptor, all suggest that it may have some distinct purpose. Still, two independently derived strains of IgD knockout mice have failed to ascribe to it a convincing immunologic role. In one strain, in which a premature stop codon was introduced into the $C_\delta 3$ domain, a subtle reduction in the total number of peripheral B cells was noted (132). In the other, which carried an insertion in its $C_\delta 1$ exon and a frameshift in $C_\delta 3$, delayed affinity maturation during T cell-dependent antigen responses was demonstrated (133).

Although not known to have any unique function, IgD, together with IgM, is a major surface component on many B cells. Because the C region genes for μ and δ are both transcribed in the same primary RNA message, differential splicing to produce either IgM or IgD is required. This particular genomic organization facilitates their coexpression, which is not possible for any other isotypes (reviewed in Chapter 5). Mature, naive B cells migrate from the bone marrow as IgM+/IgD+ cells (134) and make up about 90% of peripheral B cells in both the murine and human systems (see Chapter 6). Similarly, B cells in the primary follicles of secondary lymphoid organs coexpress IgM and IgD, but as they mature to memory cells, IgD expression is typically lost (135,136). Curiously, studies of IgM+/IgD+ splenic B cells reflect that IgD surface

expression is actually tenfold higher than IgM levels (137). This is particularly puzzling, given that δ message levels are lower than are μ mRNAs, and IgD (at least in the serum) is known to be so proteolytically labile. Perhaps helping to explain this high level of IgD expression is that fact that IgD does not need to complex with other proteins for transport to the cell surface, distinguishing it from all other immunoglobulin classes (138). It is possible that IgD's high levels of surface expression and intrinsic flexibility (139) afford it a role in the early response to antigen (123a,140).

In addition to their coexpression, IgM and IgD have a number of commonalties in terms of their function as B cell antigen receptors. Like IgM, IgD is non-covalently associated with Igα/Igβ heterodimers, which serve as the signaling component of their BCR (138). Not surprisingly then, ligation of either IgM or IgD by antigen can independently mediate activation, deletion, or anergy of B cells (141), and likewise, the signals propagated by IgM or IgD BCR seem to be the same, albeit with different kinetics (see Chapter 7). Specifically, signals transmitted through surface IgD have been reported to cause induction of APC function (142); upregulation of coreceptors B7-1 and B7-2 (143); class switching to IgM, IgG1, IgG2, IgG3, and IgA (144); and increased secretion of IgE (145). The biologic significance of many of these findings remains unclear. In fact, reports (146,147) describing a new class of germinal center IgD+ B cells (having evidence of up to 80 different somatic mutations within their V regions!) demonstrate just how little is still understood about the cells expressing—and the protein—IgD.

IgG

IgG is the predominant immunoglobulin in blood, lymph, peritoneal fluid, and cerebrospinal fluid. Collectively, it makes up more than 75% of serum immunoglobulin and is synthesized at a high rate (over 30 mg/kg/d, second only to IgA). The presence of high-affinity IgG is the hallmark of secondary humoral immune responses. Electrophoretically, IgG proteins migrate to the γ range of serum globulins, hence IgG's earlier designation as gammaglobulin. Actually, IgG is composed of four subclasses of antibody, whose salient features are summarized in Table 3. The selection of IgG subclass by a particular immune response does not appear to be random: in murine systems, anti-carbohydrate specificities tend to be IgG3, anti-protein IgG1, and anti-viral IgG2a (148,149). In man, reactivities against polysaccharide immunogens are skewed toward IgG1 and IgG2, while anti-protein and anti-viral γ antibodies are biased in the direction of IgG1, IgG3, and IgG4 (150). As an offshoot of these phenomena, clinical syndromes in which specific IgG subclasses are absent are known to present themselves as characteristic immunodeficiencies (see Chapter 43).

Perhaps the most studied feature of the IgG isotypes is their ability to activate the classical complement pathway. Although all four are capable of initiating the classical cascade, they do so to varying degrees (G3>G1>G2>G4) (150,151). Understanding the means by which the different IgG subclasses interact with specific components of complement has been difficult, complicated by many confounding reports. Results indicating that C1q is unable to bind either IgG2 or IgG4 antibodies (152) were perplexing, given that both are able to activate the classical cascade. Similarly, despite Fcγ3's higher affinity for C1q (152), IgG1 is more effective at potentiating complement-mediated cytolysis. When the site for C1q-binding was mapped to the C-terminal portion of the $C_\gamma 2$ domain near the hinge region (153,154), investigators surmised that differences in the IgG subclasses' abilities to activate complement were likely attributable to steric freedom, or lack thereof, conferred by the particular hinge of the antibody (155,156). For this reason, the longer hinge of the γ3 C region was thought to account for IgG3's higher affinity for C1q, relative to that of IgG1. Still, hinge-deletion (157) and hinge-swapping (158) experiments have yielded data that contradict the notion of the hinge being a key determinant for complement activation. Be that as it may, recall that proper glycosylation within the $C_\gamma 2$ domain is accepted as an obligatory element for fixation of complement as well.

The explanation for the difference in efficacy of lysis by complement between IgG1 and IgG3 (paradoxical, given their affinities for C1q) is even more convoluted. It may reflect that other differences between the IgG1 and IgG3 C1q sites are present that affect complement activation. Alternatively, it may derive from differences in a second, separate site in the $C_\gamma 1$ domains of these molecules which, like IgM, binds activated C3b and protects it from inhibition. This second site also likely explains the capacity of IgG2 and IgG4 to activate the classical pathway, despite an inability to bind C1q. Finally, note that IgG4 is able to efficiently recruit and activate the alternative complement cascade, distinguishing it from the other three subclasses.

Another means by which IgG antibodies communicate with the effector arms of the immune system is via the Fcγ receptors (FcγRs). A number of different IgG FcR exist (covered specifically in the section on the IgSF), each of which have their own profile and affinities for binding of the different IgG subclasses, expression patterns on different cell types, and different biologic responsibilities (159). Among the immunologic cell types implicated as

TABLE 3. *Properties of Human IgG subclasses*

Property	IgG1	IgG2	IgG3	IgG4
Disulfide linkages	2	4	5–15	2
Molecular weight	146 kD	146 kD	165 kD	146 kD
Percentage of total serum immunoglobulin	34–87%	5–56%	0.5–12%	7–12%
Average adult free serum level (mg/ml)	5.9 ± 2.6	3 ± 2.5	0.6 ± 0.55	0.9 ± 0.25
Macrophage binding by FcγR	+	–	+	+
Placental transfer	+	++	+	++
ADCC	+++	+	+++	+
Classical complement activation	+++	++	++++	+
Alternative complement activation	+	+	+	+++

Adapted from Simard and Mak (140).

important binders of IgG are macrophages, polymononuclear cells, and lymphocytes (including B cells). Interactions with these receptors cause many functional effects, including phagocytosis (160) and antibody-dependent cell-mediated cytotoxicity (161), both of which ultimately lead to the destruction of the bound antigen. Specifically, the hierarchy for ADCC by mononuclear cells is IgG1, IgG3 > IgG2, IgG4 (152,162). Signals transmitted via FcγR also modulate lymphocyte function by means of up-regulation or down-regulation of antigen presentation, cytokine release, cytokine receptor expression and/or sensitivity, and even immunoglobulin secretion. Even soluble FcγR are known to bind IgG, although the significance of this finding is unclear (163). Finally, IgG FcR also permit transplacental movement of maternal antibodies during gestation (164). This provides the developing fetus with a source of high-affinity serum immunoglobulin that is able to interact with complement to mediate biologic effects at a time at which it has no other form of specific humoral immunity. It should not be overlooked that IgG molecules are the most stable isotype in serum (with a half-life of over 3 weeks), further maximizing their utility in this endeavor—even into the post-natal period.

Binding of the four IgG subclasses by the different FcγR varies in terms of the specific contact residues involved for each respective ligand–receptor pair. Generally, although the IgG binding sites for the FcR are thought to largely overlap, the precise elements responsible for interaction likely have subtle differences. By consensus, research into these issues has suggested that the sites are bipartite, consisting of a site on the C-terminal portion of the hinge and also reliant upon residues found in the portion of the $C_\gamma 2$ domain already implicated in C1q-binding (86,165). Because the four IgG isotypes differ considerably in these regions, this would fit nicely with their noted differential binding of the varied FcγR.

IgA

IgA is the major immunoglobulin in external secretions such as saliva, mucus, sweat, gastric fluid, and tears. Moreover, it is also the major immunoglobulin of colostrum and breast milk, where it provides the neonate with a readily available source of intestinal protection against pathogens (167). The secretory forms of IgA are exclusively polymeric, including J chain and SC in the manner described previously. In addition, IgA—present predominantly in its monomeric form—is also an important component of serum Ig, where it makes up 10% to 15% of the total. The synthetic rate of IgA is roughly double that of IgG, such that total daily IgA production outpaces that of all other immunoglobulins combined. The majority of IgA synthesized is in the secretory form, with the largest fraction of IgA plasma cells residing in the subepithelial mucosa of the small intestine. Because secretory IgA coats all external surfaces except skin, it is rightly considered a first line of defense against organisms that would invade via mucosal routes. IgA's role in mucosal immunity (see Chapter 27) is phenotypically evident in persons with the most common genetic defect of the humoral immune system, IgA deficiency (see Chapter 43). Individuals with this condition are susceptible to invasion across mucosal barriers and typically present clinically with recurrent infections of this type.

Serum and secreted IgA originate from separate pools of B lymphocytes: plasma cells in specialized sites of the respiratory, urogenital, gastrointestinal, and mammary tissues produce the IgA found in secretions, while the IgA in serum emanates from plasma cells in the bone marrow, lymph nodes, and spleen. Despite this compartmentalization of production, antigenic exposure occurring at either mucosal or systemic sites will prime the development of both secretory and serum IgA responses simultaneously (168).

In humans, the two IgA subclasses, IgA1 and IgA2, show an interesting division of expression which affects their resultant biologic utilities. IgA1 exists primarily as a monomeric molecule, and accordingly is the main IgA isotype in plasma (refer to Fig. 20). In bone marrow, about 90% of IgA-secreting plasma cells make IgA1 (169). IgA2, on the other hand, is usually found as a polymer. Recall that the main structural difference between these two isotypes is localized to the hinge. Whereas the IgA1 subclass has a higher concentration of carbohydrate in its hinge region (protecting it from most forms of proteolytic degradation), the IgA2 isotype has deleted much of that hinge region—presumably as an evolutionary response to bacterial IgA1-specific proteases (170). Thus, it is consistent that the broadly protease-resistant form (IgA1) should predominate in serum to maximize its lifespan, while the targeted protease-resistant subclass (IgA2) should prevail where bacterial exposure is more common.

IgA does not efficiently induce inflammatory responses. Rather, it is believed to protect primarily by exclusion, binding and cross-linking pathogens to prevent their uptake across epithelia and facilitating their expulsion in mucus excretions (123a,140). It is noteworthy that inflammatory responses localized to mucosa would likely be detrimental to barrier function, as tissue damage could compromise the integrity of epithelial surfaces. While IgA does have the ability to fix complement via the alternate cascade, this ability is restricted to IgA1. IgA can also opsonize antigens for phagocytosis; this is accomplished via a specific Fcα receptor (FcαR) found on macrophages, monocytes, and neutrophils. This provides a mechanism for IgA immune complexes that accumulate at mucosal surfaces to be engulfed and processed. The FcαR is known to bind secretory IgA with higher affinity than serum IgA, but strangely, the site on IgA that is recognized seems to be unrelated to the J chain or SC which distinguish the secretory and serum forms (66). Rather, in a manner unique from all other IgSF FcR, which see a hinge-proximal site in the C_H2 (or equivalent) domain, the FcαR sees a site bridging the domain boundary between $C_\alpha 2$ and $C_\alpha 3$, reminiscent of SPA binding (66). Finally, IgA has also been shown to induce eosinophil degranulation via the FcαR, implicating it in antiparasite immunity. Given that many parasites gain access to host tissues by crossing mucosal barriers, this is a logical biologic activity for IgA as well.

IgE

IgE is present in serum in the lowest concentration of all the immunoglobulins. Its rate of synthesis is between 25- and 2,000-fold less than each of the other isotypes, it has the shortest serum half-life, is unable to activate either the classical or alternative complement cascades, and lacks the ability to opsonize antigens. Nonetheless, IgE's biological effects more than compensate for these shortcomings, due to the profound efficiency of its behavior. The principle function of IgE is to arm basophils and mast cells with specific antigen receptors. These cells in turn act as potent dispensers of inflammatory reactions (see Chapter 32).

Plasma cells that produce IgE are chiefly found in the lung and skin. Upon its release from these B cells, circulating IgE is quickly bound by a high-affinity Fcε receptor (FcεRI; $K_D = 10^{-10}$ M) found

on these granulocytes, allowing the IgE molecules to stably remain on the cells for weeks or months. Once primed with many such antigen-specific receptors (recognize that cells bearing FcεRI can have IgE molecules of many different reactivities on their surfaces, unlike the case for antigen-specific B cells), multivalent antigen can then cross-link the bound IgE, indirectly cross-linking the FcεRI molecules as well. Ultimately, this causes mast cells and basophils to release granules containing inflammation-mediating substances and chemoattractants for a variety of cell types. The granule contents of mast cells and basophils are powerful, able to induce rapid responses—including mucous secretion, coughing and sneezing, vomiting, diarrhea, and inflammation. While this type of response can be vital in the clearance of parasites (see Chapter 38), it has the unfortunate consequences of also causing allergy (see Chapter 35) and anaphylaxis in predisposed individuals. In such atopic individuals, it has been seen that increased amounts of IgE are synthesized and found on the surfaces of mast cells and basophils, likely explaining their predilection for these inappropriate responses.

Other cells types also express the high-affinity FcεRI, including Langerhan's cells (171,172) and eosinophils (173), but the rationale for its presence there is yet to be definitively elucidated. In addition, the CD23 surface antigen has also been shown to be a low-affinity IgE receptor (designated FcεRII). Among other cell types, CD23 is known to be expressed on monocytes and some follicular B cells. In fact, monocytes can even be induced to secrete a soluble form of FcεRII (174), but once again the significance of this finding is unclear. Considering the low levels of circulating IgE, the relatively low affinity of the receptor, and the fact that CD23 is known to interact with CD11/CD18, there is doubt as to whether IgE is even an important ligand for this receptor *in vivo*.

Like the pIgR and several of the FcγRs, the FcεRI molecule is also a member of the IgSF (detailed further in the following section). Interaction between IgE and its high-affinity receptor was the first well-characterized Ig-FcR ligand–receptor pair of this type. Originally, studies using synthetic peptides as specific inhibitors of IgE–FcεRI binding identified a 76-amino acid polypeptide spanning $C_\varepsilon2$–$C_\varepsilon3$ as the FcR recognition site on IgE (175). Subsequently, this localization was refined further to a site in $C_\varepsilon3$ analogous to the FcγR site on $C_\gamma2$. Unlike the binding situation for IgG that was also dependent upon residues in the hinge regions (see above discussion), the extra $C_\varepsilon2$ hinge domain of IgE is not believed to play a significant role in the interaction (176).

THE IMMUNOGLOBULIN SUPERFAMILY

Evolution of the Immunoglobulin Superfamily

Soon after the sequencing and structural analyses of antibodies revealed the protein motif of the immunoglobulin domain (7,177), it became apparent that evolution had incorporated Ig homology domains in a variety of other important molecules as well. The sequencing of MHC genes, TCRs, and the pIgR, among others, demonstrated the use of both V region- and C region-type domains by a number of cell-surface proteins of the immune system. Contemporaneously, a number of cell adhesion molecules (CAMs) involved with neurite outgrowth in developing axons were also found to contain Ig-like domains (reviewed in ref. 178). It was quickly recognized that a large family of genes that contained putative immunoglobulin folds existed (2,3,179), whose members were globally implicated in issues of molecular recognition and/or cellular adhesion. Comprised of several multigene families in their own right (V_H, V_L, TCRα, TCRβ, TCRγ, TCRδ, MHC I, MHC II, Sialoadhesin, CAM, etc.), the term *immunoglobulin superfamily* was adopted to refer to this diverse group of genes, which each contained one or more Ig homology domains.

Currently the IgSF encompasses well over 100 genes, and extends across several phylogenetic boundaries (reviewed comprehensively in ref. 4). Disparate species in which IgSF members have been identified include chicken, zebrafish, tunicates, grasshoppers, squid, *C. elegans*, sponges, and *S. cerevisiae*. In addition, reports identifying proteins containing structures reminiscent of immunoglobulin homology domains from prokaryotic organisms (180,181) raise the possibility that this archetypal structure antedates even eukaryote evolution. The discovery of new molecules with novel functional attributes (for this class of proteins) also continues to expand the role of IgSF members. For instance, while the preponderance of immunoglobulin homology domain-containing proteins that have been identified are either cell surface or secreted proteins involved in recognition and adhesion events, a newer class of intracellular muscle proteins (titin, telokin, etc.) belonging to the IgSF demonstrate that the immunoglobulin fold structure is adaptable to an assortment of functional capacities.

The evolution of the IgSF has been the subject of considerable scientific speculation and potentially has implications for both the development of the vertebrate immune system and the process of organogenesis in general. Based upon the overwhelming number of IgSF members that possess adhesive qualities, it has been proposed that the first IgSF molecules were simply single Ig domain extracellular proteins that served as primordial "cellular glues" (2,3). Substantiating this argument is the noted stability of the compact β-barrel structure of the immunoglobulin fold, which would foster its utility in harsh extracellular environments. Further bolstering this hypothesis is the fact that numerous IgSF proteins participate in both homotypic and heterotypic interactions with other IgSF molecules, demonstrating their potential to act as adhesion molecules.

Some evidence suggests that IgSF proteins have promoted clustering of cells since the earliest stages of eukaryotic development. For example, the yeast *S. cerevisiae* uses the IgSF glycoprotein α-agglutinin to mediate cell–cell contact during mating (182). IgSF forebears may have also allowed the first examples of rudimentary organogenesis in phylogeny. The slime mold *Dictyostelium*, which bridges the gap between unicellular and multicellular eukaryotes, uses a protein possessing a region with striking similarity to an immunoglobulin domain for the purpose of forming aggregations called "fruiting bodies" when conditions are nutrient-scarce (183,184).

Finally, it has also been put forth that ancestral IgSF glycoproteins may have mediated the first evolutionary examples of allorecognition in colonial invertebrates (185). In defense of this proposition, it is noteworthy that two examples of metazoan receptor tyrosine kinases with purported recognition functions have been shown to possess extracellular Ig-like domains, one from the cnidarian *Hydra vulgaris* and the other from the marine sponge *Geodia cydonium*. From these data, then, it is possible to make tentative, yet tenable, conjecture that the complex cellular and molecular interactions of the vertebrate immune system (mediated in no small part by members of the IgSF) may in fact be an outgrowth of this primitive allorecognition. In this light, the notable analogy between vertebrate graft acceptance/rejection reactions and colonial invertebrate fusion/rejection phenomena perhaps takes on new significance (see Chapter 18).

Regardless of its derivation, the immunoglobulin domain has obviously proven to be a pliable evolutionary substrate, amenable to mutation and diversification for a number of important reasons. First, as was noted for the actual domains of immunoglobulins, the primary structure of these units can vary dramatically without appreciably altering their tertiary structure (186,187). This is particularly evident in the interconnecting loops that join the β strands, allowing them to diverge rapidly to perform a multitude of distinct functions. Second, most Ig domains are encoded by discrete exons, facilitating their duplication by relatively simple genetic events. This one-domain-per-exon rule is also conducive to alternative splicing phenomena, encouraging differential expression of IgSF molecules as well. This is further accommodated by a splicing convention followed by most IgSF exons: The 3′ end of one exon is always the first position of a codon, while the 5′ end of the next tandem unit begins with the second position of a codon. Thus, immunoglobulin homology domains of IgSF proteins may be easily duplicated in tandem (the *C. elegans* muscle protein twitchin contains 26 Ig-like domains) and shuffled to create new genes with the capacity to diversify both somatically and evolutionarily. Finally, the propensity of Ig domains to form homotypic and/or heterotypic dimers (also demonstrated by immunoglobulin proper) forms the basis for proteins which serve as receptor and ligand molecules. These combinatorial associations enhance their diversification potentials still further.

Despite the inherent complexity in a gene superfamily containing such vast disparities in its members' functional qualities, it was recognized early on that IgSF proteins could be subdivided into distinct "sets" on the basis of sequence and structural analyses (2). These groupings are based upon the arrangement of the β strands of the immunoglobulin fold and are schematized in Fig. 21. Note that while V-set domains are composed of a four-strand sheet and a five-strand sheet (as detailed earlier), C-set domains have four-strand and three-strand layers; these are discriminated on the basis of placement of the D strand in the sheet of strands A, B, and E (the C1 set) or with the layer formed by strands G, F, and C (the C2 set). However, studies with the IgSF muscle protein telokin have revealed a new "I set," which has domain features that are intermediary between the V and C1 sets (188). Moreover, many IgSF adhesion molecules and cell-surface receptors likely belong to this I set, rather than to the sets to which they were previously ascribed. In any case, the IgSF remains a fascinating collection of proteins with structural similarities but a wide array of functional abilities. While immunoglobulin remains the definitive example of this class of molecules, a number of IgSF proteins are also of particular significance to humoral immune responses, and their structural and functional characteristics are briefly summarized in the following sections.

Fc Receptor Molecules

FcR allow antibodies to interact with cells of both the specific and non-specific immune systems. In so doing, FcR connect humoral immune responses to cellular immune responses, and more globally, acquired immunity to that of innate immunity. These contacts play two vital roles in the biology of immune functioning. First, FcR allow antibodies to act as "flags" signaling the need for certain cellular effector events, such as phagocytosis and ADCC. Second, the different FcR facilitate antibody acting as a mediator of overall immune regulation. The signals they transmit can induce changes in cytokine secretion, expression of cell-surface receptors, and extensive differentiation programs (189).

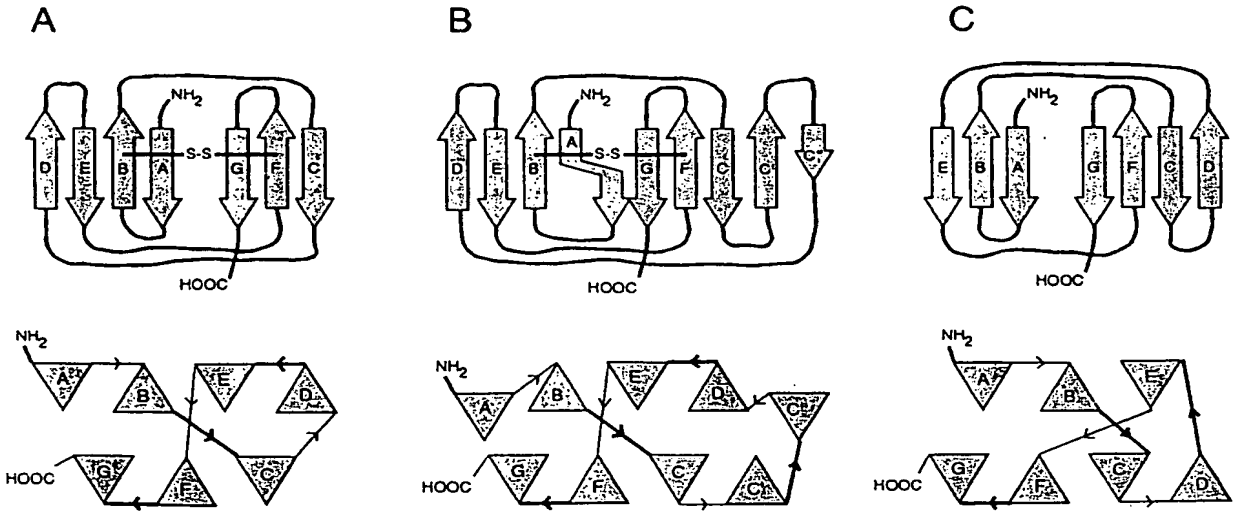


FIG. 21. Topology of different immunoglobulin domain types. Diagrams of the (**A**) C1 set, (**B**) V set, and (**C**) C2 set are presented. In the upper part of the figure, β strands are depicted as *broad arrows* and their intervening loops by *thin lines*. Note that the V-type domains have five- and four-stranded faces, while C1- and C2-type domains have four- and three-strand faces. The C region-like structures are discriminated on the basis of placement of their D strand. In the lower part of the figure, an end-on view of the different β-barrels is shown. *Triangles* (*with their apex at the top*) symbolize β strands running out of the plane of the paper; *triangles* (*whose apex points down*) are β strands traveling into the paper. *Bold lines* represent connecting loops at the top of the immunoglobulin fold; *thin lines* indicate connections at the bottom of the domain. (From ref. 4, with permission.)

Three large classes of molecules can bind Fc regions: glycosyltransferases, which recognize oligosaccharide derivatives on antibodies, lectin-like molecules, and receptors belonging to the IgSF. Of the "true" FcR that recognize antibody protein determinants rather than carbohydrate, all FcR thus far identified belong to the IgSF, other than the low-affinity IgE receptor (CD23/FcεRII). These molecules include FcγR I, II, and III (CD64, CD32, and CD16), FcεRI, FcαR (CD89), and the pIgR, which has already been discussed. All cells of lymphoid origin express FcRs, although the profiles and isotype specificities between lineages can vary greatly (reviewed in refs. 190 and 191). While receptors for all classes of immunoglobulin have been described as biological activities, human FcμR and FcδR have not yet been cloned. Thus far, the FcγR proteins and FcεRI are the most well characterized examples of these molecules.

FcγR Molecules

Receptors for the Fc portion of IgG are of three types (reviewed in ref. 192). FcγRI (CD64) is a high-affinity receptor and the only one able to bind monomeric IgG. It possesses three extracellular Ig-like domains. FcγRII (CD32) and FcγRIII (CD16) are both low-affinity receptors that bind IgG-containing immune complexes. They each have only two extracellular Ig homology domains. Schematic diagrams of the FcγRI, FcγRII, FcγRIII complexes, along with the FcεRI and FcαR complexes, are presented in Fig. 22.

FcγRI is a 70-kD glycoprotein that is constitutively expressed at low levels on monocytes and macrophages. IFN-γ upregulates its levels on these cells, and also can induce its expression by neutrophils. FcγRI's affinity for IgG is highest for the IgG1 and IgG3 subclasses ($K_D = 10^{-8}$ M), tenfold lower for IgG4, and will not bind IgG2. Functionally, the primary effect of cross-linking FcγRI molecules appears to be the potentiation of both ADCC and phagocytosis. As IFN-γ enhances both of these activities by the cell types known to express FcγRI, this would fit well with their being important roles for the receptor.

Like surface immunoglobulin, FcγRI requires accessory proteins in order to transmit signals. This is, in fact, a common feature of most FcRs (except for FcγRII) and an interesting parallel between the IgSF antigen receptors (BCR and TCR) and the IgSF "indirect" antigen receptors (the FcR), which use antibody to bridge the span between FcR and antigen (193). In the specific case of FcγRI, the actual signaling molecule is a 12γ-kD transmembrane protein designated the "γ-subunit" or, more generally, FcRγ. This nomenclature can be particularly confusing, as the "γ" of FcRγ refers not to the fact that it is part of the of FcγR complex (the receptor for γ-class immunoglobulin), but rather to γ as an individual subunit of a multi-molecule complex. In any case, for the FcγRI *complex*, the α subunit is the actual IgSF protein FcγRI, and the γ subunit is FcRγ, which forms a disulfide-linked homodimer. Complicating terminology further, FcRγ is also a subunit of other FcR complexes, including that of the FcγRIIIA and those of the non-FcγR, FcεRI and FcαR. Intriguingly, FcRγ is a close homologue of the TCR-associated protein CD3ζ. In fact, CD3ζ cannot only heterodimerize with FcRγ, but also has been shown to be capable of functionally substituting for FcRγ as the signal-transducing subunit of the FcγRIIIA receptor complex (194).

The situation for FcγRII (CD32) is even more complex. FcγRII is the product of three distinct but homologous genes: FcγRIIA, FcγRIIB, and FcγRIIC. This is further complicated by the fact that at least two of the FcγRII genes are alternatively spliced to generate multiple isoforms (195). The FcγRIIA gene gives rise to two transcripts: FcγRIIa1, which has a transmembrane domain, and FcγRIIa2, which lacks it. The FcγRIIB gene has three isoforms—FcγRIIb1, FcγRIIb2, and FcγRIIb3—generated by differential splicing and alternative polyadenylation processing. Collectively, the FcγRII variants are the most ubiquitously expressed FcγRs, being present on monocytes, macrophages, neutrophils, B lymphocytes, megakaryocytes, and platelets. Specifically, megakaryocytes express FcγRIIA (both isoforms), B lymphocytes express FcγRIIB (b1 and b2 transcripts) and FcγRIIC, and cells of myelomonocyte derivation produce at least one or more isoforms from all three genes (195).

Functionally, due to their expression on many cell types, FcγRII signals cause diverse effects. When cell surface FcγRII engage IgG immune complexes (all subclasses, with varying affinities), they potentiate several biologic changes, most immunoregulatory in nature. Generally, these signals down modulate IgG-, IgA-, and IgE-mediated activations of a number of cell types, including monocytes and macrophages, granulocytes, mast cells, and Langerhans and other dendritic cells. They also induce platelet aggregation at the site of immune complexes and effect B cell feedback inhibition by down-

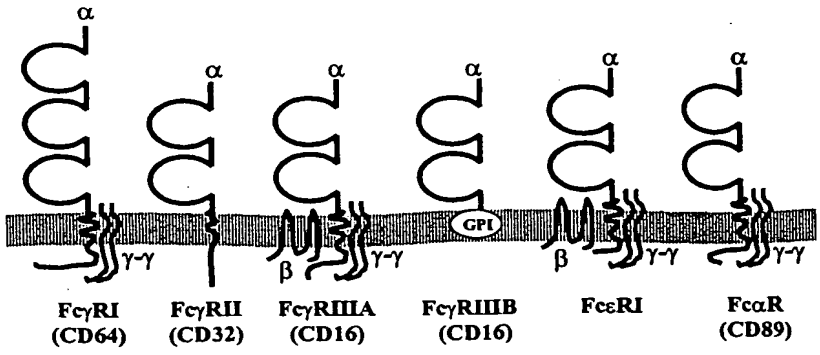


FIG. 22. Schematic diagram of human immunoglobulin Fc receptors belonging to the IgSF. Each Ig domain is depicted as a *rounded bulge*. The α chains are the components of the receptor complex that determine binding specificity. β and γ chains are responsible for association and signal propagation by the receptor(s).

regulating both proliferation and antibody production. Specifically, FcγRIIB are known to suppress BCR-mediated activation signals when the two are coaggregated (196). The signal for this inhibition is brought about in a manner unique among all IgSF FcR. As single-chain receptors without accessory proteins, FcγRII are able to transduce their own signals. They do so by way of an immunoreceptor tyrosine-based inhibition motif (ITIM) present in the cytoplasmic region of the protein (197).

Specific binding of IgG by FcγRII was initially mapped to the C-terminal portion of the second extracellular Ig-like domain of the receptor (198). This primary site, comprising residues Asn 154–Ser 161, has since been revised to also include domain 2 stretches Ser 109–Val 116 and Phe 129–Thr 135, along with domain 1 contacts (199). In sum, a three-dimensional model of the entire FcγRII extracellular region predicts that loops of both Ig-like domains that co-localize to the domain interface are responsible for the recognition of IgG.

The final IgG FcR is the FcγRIII (CD16). FcγRIII has two extracellular Ig-like domains and is encoded by two separate genes whose expression varies by cell type. On monocytes, macrophages, and NK cells, it is a transmembrane glycoprotein called FcγRIIIA. The FcγRIIIA receptor protein has three accessory proteins with which it is complexed. The first is a 30-kD "β-subunit" having four transmembrane regions, which is also a component of the FcεRI complex. The other protein(s) associated with FcγRIIIA is a homodimer of FcRγ subunits, which, as explained earlier, is also part of the receptor complexes of FcγRI, FcεRI, and FcαR (200). The other FcγRIII gene encodes a glycophosphoinositol-linked protein termed FcγRIIIB that is expressed on neutrophils. Individuals deficient in this gene suffer from a condition called paroxysmal nocturnal hemoglobinuria, characterized by increased susceptibility to infection and delayed clearance of immune complexes (201). FcγRIII's binding preference is for IgG1 and IgG3, both of which it binds with low affinity only in the form of immune complexes. Biologic activities fostered by FcγRIII include ADCC, phagocytosis, and transport of internalized Ab-Ag complexes to the antigen-presentation pathway.

FcεRI and FcαR Molecules

The high-affinity IgE receptor (FcεRI) was the first and best characterized of the FcR (202). FcεRI is a transmembrane protein having two extracellular Ig homology domains and, like previous examples, associates with accessory proteins for signal-transmission purposes. It is expressed on mast cells, basophils, eosinophils, Langerhans cells, and on the monocytes of atopic individuals (203). The proteins associated with FcεRI are the same as for FcγRIIIA: one β subunit and a homodimer of FcRγ proteins. Gene targeting experiments to verify the roles these proteins play in the FcεRI complex have yielded reassuring results: homozygous deletion of the IgSF chain of the receptor created animals that were predictably resistant to anaphylaxis (204); disruption of FcRγ caused the same phenotype, plus defects in ADCC and phagocytosis consistent with the γ-subunit's participation in other receptor complexes (205,206).

The Fcε-binding site on FcεRI has been localized to three regions of the second extracellular Ig domain (198). It is important to remember that serum IgE binds to the receptor in a high-affinity interaction not dependent on antigen (unlike all other antibody isotypes, which must first bind antigen in order to be recognized by their respective FcR). This allows polyvalent immunogens to bind effector cells directly, without the need for conformational change of the immunoglobulin and/or immune complex formation. This contributes to the rapidity of the response exhibited by cells expressing FcεRI. The specific biologic effects of cross-linking the receptor were discussed previously in the section on IgE function.

The final FcR of the IgSF to be covered is that for IgA. The FcαR (CD89) is the most recently identified and least characterized of the different Ig receptor classes. It possesses two extracellular Ig domains and is expressed by monocytes, macrophages, neutrophils, and eosinophils. Several isoforms have been identified: a cell surface FcαRa form, which has intracellular and transmembrane domains, an FcαRb form lacking these domains that is both secreted and associated with the cell surface (by an unknown mechanism), and even an isoform lacking the membrane-proximal Ig-like domain. Structurally, the FcαR has homology with the FcγR molecules. Like FcγRI, FcγRIII, and FcεRI, the IgA receptor complex includes the FcRγ homodimer as a signaling component. The particulars of the receptor binding site on the FcαR are not yet determined. The site it recognizes on IgA and the effects mediated by FcαR-binding were detailed in the section on IgA function.

Coreceptor CD4 and CD8 Molecules

Antigen-recognition functions in the body are not limited to immunoglobulin but are also performed by receptors on T cells (TCR), which bind antigenic peptides. In a defining event of immune responses, these antigenic fragments are "presented" to T cells within the context of molecules of the MHC. Recognition of MHC/antigen by TCR is the fundamental biologic interaction responsible for initiation, perpetuation, and mediation of cellular immunity. As pertains to antibody, binding of MHC/peptide complexes by TCR is also vital for recruitment of T cell help needed in many humoral immune responses. While the details and effects of these vital interactions are well beyond the scope of this chapter (see Chapters 8–13), many IgSF proteins play key roles in assuring its productive functional outcome. Of these, the TCR "coreceptors" CD4 and CD8 are crucial components worthy of mention here.

CD4 and CD8 were among the first non-immunoglobulin IgSF members for which structural information became available (reviewed in ref. 207). These molecules are each expressed on the surface of T cells, where they participate in TCR/MHC interactions (schematized in Fig. 23) by engaging non-polymorphic regions of the MHC in low-affinity interactions (208,209). T cells break down into two major functional subclasses—helper T cells and cytotoxic T cells—characterized by different responses to antigen. Both T cell types utilize the same group of TCR genetic elements to compose their specific antigen-receptors, however. Ordinarily, although there are notable exceptions, T cell effector functions correlate with the type of MHC protein with which they interact. MHC I molecules specify cytotoxic T cell responses and are found on most cell types of the body. MHC II molecules, on the other hand, dictate helper T cell functioning and are more restricted in their expression, typically found only on "professional" antigen-presenting cells. CD8 molecules bind to class I MHC proteins, while CD4 molecules mediate interaction with class II; thus, these two proteins play an important role in determining what type of response a particular T cell is likely to mediate.

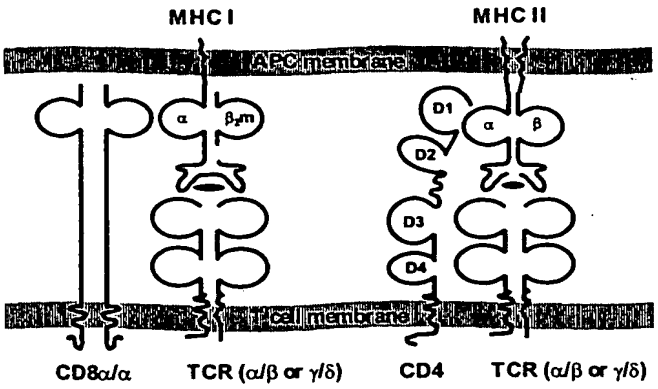


FIG. 23. Schematic diagram of the CD8 and CD4 coreceptor molecules. The figure shows both TCR/MHC coreceptor complexes on the same membrane, although T cells express either CD8 or CD4 for the majority of their lifetime. *Bulges* represent Ig domains, and *gray ovals* signify peptide presented by MHC molecules. Only the CD8α/α homodimer is schematized here, although the CD8α/β heterodimer presumably binds MHC I in similar fashion. The models demonstrate the simplest stoichiometry for association; other possible modes of interaction are discussed in the text.

Both CD8 and CD4 are glycoproteins and, in their most common incarnations, possess transmembrane segments and short cytoplasmic tails (210). The intracytoplasmic regions of both molecules interact with the src-like tyrosine kinase $p56^{lck}$ (211), which presumably serves to allow signal-transduction necessary for proper thymic selection (212) and activation (213) of T cells. Reinforcing this idea is the fact that co-ligation of CD4 or CD8 with TCR greatly enhances stimulation of T cells relative to that of cross-linking TCR alone (see Chapters 12 and 13). In addition, CD8 and CD4 also increase the avidity of interaction between TCR and MHC, by virtue of their action as adhesion molecules between the two cell membranes. Collectively, these behaviors have been estimated to boost antigen recognition over 100-fold from that of basal levels (TCR engagement by MHC/Ag only). However, despite these similarities in their biologic effects, structurally CD8 and CD4 have many important differences. While many details are still unresolved, a number of crystals involving these two proteins have been solved by x-ray diffraction, permitting a thorough examination of their salient characteristics.

CD8

CD8 exists as a disulfide-linked dimer in one of two forms. A homodimer of two CD8α subunits was the first human isoform identified (214). Subsequently, a CD8α/β heterodimer was described as well (215,216). Both proteins are 34-kD and have homologous (although only 17% identical) N-terminal Ig-like domains, extended hinge regions of 50 (α) and 30 (β) amino acid residues, single transmembrane stretches, and short cytoplasmic tails. While the CD8β chain lacks residues necessary for interaction with $p56^{lck}$, the heterodimer is still capable of interacting with it via the α subunit. Nevertheless, several lines of evidence indicate that a specific role for the heterodimer may exist, apart from that of the homodimer. Differences in functional effects (217), thymic selection (218), avidity for MHC (219), and $p56^{lck}$ activity (220) have all been attributed to the CD8β chain. In addition, there are differences in expression of the different isoforms: thymocytes and peripheral T cells are CD8α/β+, TCR γ/δ+ intraepithelial lymphocytes (IEL) are primarily CD8α/α+, and TCR α/β+ IEL express either of the two molecules. Finally, investigations into CD8β have divulged that there are actually two genes encoding this protein (221) (i.e., the locus has been recently duplicated), which—along with alternative splicing phenomena—results in as many as seven unique CD8β isoforms being expressed (222). Four of these transcripts lack the transmembrane region of the message, raising the possibility that some forms of CD8β may be secreted. In sum, the story of this protein subunit, and of the CD8 heterodimer that derives from it, is still an active area of research that is yet to be clarified.

The enumeration of the CD8α/α homodimer, thanks in large part to two crystal structures, is somewhat further along. The primary advance in this regard was the solving of the amino-terminal domains of human CD8α/α (223). This study revealed that these 113-amino acid segments formed V-type Ig domains consisting of four- and five-strand layers (see Fig. 24). In agreement with its V region-type topology, CD8α domains were shown to dimerize with one another via their five-strand faces, as do immunoglobulin V domains (see Colorplate 7). Two significant structural disparities between CD8α and immunoglobulin V regions were also recognized. First, the C'-C" loop (corresponding to CDR2) is extended in the α subunit (note the right side of Fig. 24D). Second, while the usual intradomain disulfide bridge between β strands B and F was identified, an unpaired cysteine in strand C is also conserved. In rodent CD8α this residue has been shown to form the intradomain cystine together with the Cys of strand B (224). The α subunit hinge region is extensively glycosylated (via O-linkages), and this is thought to promote its adopting an extended structure (223,225–227). This is particularly important because an elongated conformation of the hinge would be necessary to allow the N-terminal Ig domain to interact with the MHC I molecule appropriately (refer back to Fig. 23).

The other important structural features of CD8 concern its interaction with class I MHC molecules. Initial experiments into these questions indicated that the CDR-like loops of CD8α were involved with recognizing a negatively-charged region on the α3 domain of MHC I (228). The aforementioned crystal structure supported this conclusion by demonstrating that these same loops were the only region on CD8 where positive charge was localized. Mutational studies performed after the crystal was solved also implicated residues in the A and B strands of the CD8α protein as contact points with the α2 domain of MHC I (229). As each CD8 homodimer has two α chains, and as the A and B strands of each chain are not on the dimerizing face of the subunit, this implied that CD8's interaction with MHC I could be bivalent (i.e., one CD8α/α and two MHC class I proteins). The publication of a crystal structure of the complex containing CD8α/α and MHC I plus peptide has seemingly resolved these issues (230). The homodimer was shown to have contacts with not only the α3 domain, but also the α2 domain, and even with the β2-microglobulin subunit of MHC class I. Strikingly, the negatively-charged region of the α3 MHC I protein fits between the CDR-like loops of the two CD8α subunits in the fashion of classical antibody–antigen interactions! However, because both CD8 subunits are needed for the binding of one such loop, it would appear that the stoichiometry of the CD8/MHC I interaction is in fact 1:1. Because the clustering of receptor complexes is likely an important feature for the generation of intracellular signals, this is a crucial piece of information, as shall be seen for CD4.

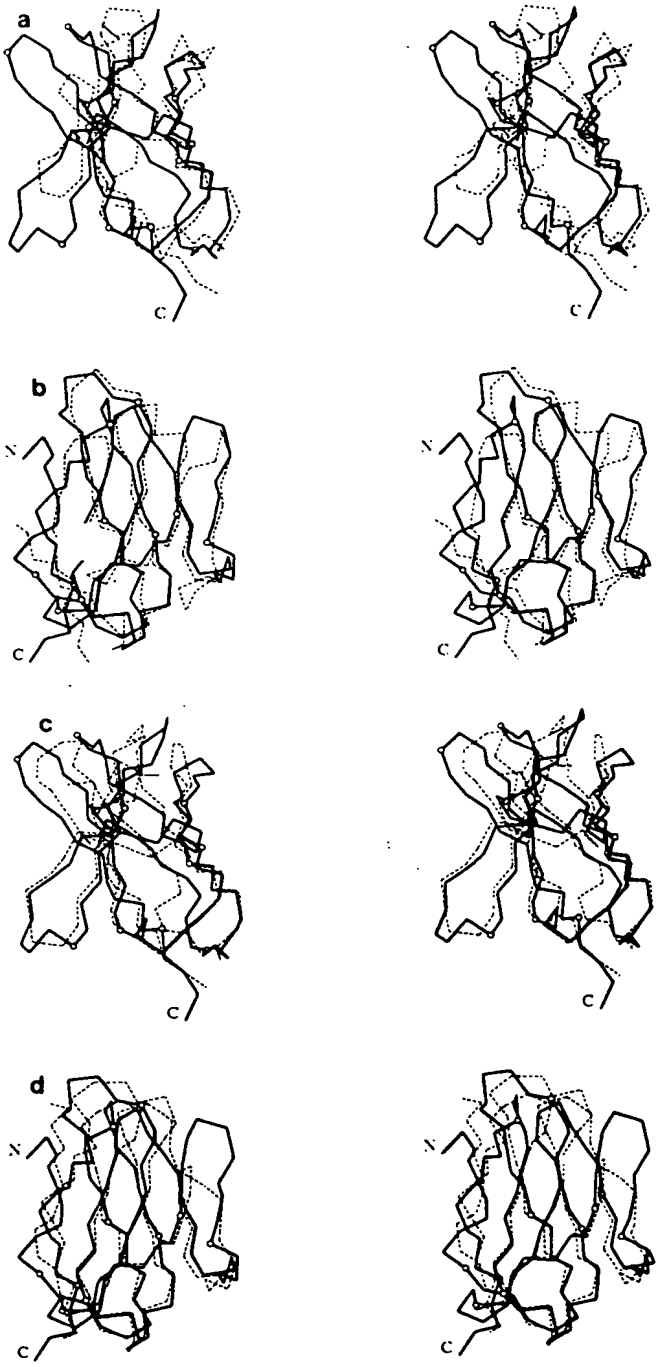


FIG. 24. Stereoviews of the α-carbon backbones of CD8α, CD4 domain 1, and the V_L of the antibody REI. In (**a**,**b**) the Ig domain of CD8α (*solid lines*) is superimposed on domain 1 of CD4 (*dashed lines*). In (**c**,**d**) V_L (now in *dashed lines*) is overlaid by CD8α. Parts (**a**) and (**c**) are side views (parallel to the dimerization surface); (**b**) and (**d**) are perpendicular to the β-sheet faces. In all cases, the CDR loops are at the top of the figure. Comparing parts (**a**) to (**c**) illustrates the truncation of the F–G (*top left*) and C–C′ (*bottom left*) loops of CD4 D1 relative to CD8α and V_L. Comparison of the N-termini [left edge of (**b**) and (**d**)] shows the shortening of CD4 D1's A strand as well. The CDR2-like C′-C" loop [upper right of (**b**) and (**d**)] demonstrates that this segment is elongated in both CD8α and CD4 D1 relative to V_L. (From ref. 207, with permission.)

CD4

CD4 has four extracellular Ig homology domains (D1–D4) and is thought to exist as a 55-kD monomer on the cell surface (231). Scientifically, CD4 became the center of intense scrutiny when it was demonstrated to be the molecule utilized by HIV for attachment to T cells. Like immunoglobulin, proteolytic analyses established that CD4 generated stable fragments upon cleavage. These, in turn, were the initial substrates for crystallographic study. The amino-terminal (and T cell membrane-distal) D1D2 segments were the first regions of CD4 structurally determined, as they had been shown to contain the HIV-binding site. These studies (232,233) described an N-terminal V-type Ig domain (D1) and a smaller, unusual Ig-like domain (D2), each with features unique among previously reported IgSF structures.

D1 is a four- and five-strand domain (see Colorplate 7) that maintains the normal core intradomain disulfide bond and its associated residues. However, by comparison with immunoglobulin V regions, it became apparent that part of the A strand was missing (see Fig. 24). Similarly, the two loops connecting β strands C to C′ and F to G were both shortened in length. Since amino acids in these positions of immunoglobulin (and CD8) participate in dimerization events, this was taken to be reflective of the fact that CD4 was not known to dimerize. Like CD8, the CDR2-homologous C′-C" loop of CD4 D1 is also extended relative to immunoglobulin (compare the right edges of Figs. 24B and 24D). In fact, a Phe residue found on this lengthened segment has been shown to be crucial for binding of HIV gp120 to CD4. Another interesting characteristic of these crystals concerns the D1 to D2 domain connection. Note in Colorplate 7 that the G strand of D1 is contigous with D2's A strand (contrast with the elbow peptides connecting V and C domains). As a result, D1 and D2 have a large amount of longitudinal contact and, accordingly, little flexibility to move relative to one another.

The D2 domain is even more peculiar. D2 is smaller than most Ig domains, and it consists of only seven β strands, like Ig C-set domains (see Colorplate 7). Unlike the canonical C-type domain, D2 has switched the placement of one of its strands from one face to the other (i.e., it belongs to the C2 set; refer back to Fig. 21). Remarkably, D2 also fails to conserve the core residues necessary for the typical intradomain cystine. Curiously, its intradomain disulfide linkage is between cysteines found in the same β-stranded sheet (i.e., its disulfide bond is *intra*sheet instead of the usual *inter*sheet). Despite this idiosyncratic arrangement, D2 forms an Ig fold consistent with other Ig domains—a powerful testimony to the principle of tertiary structural conservation in the face of primary structural variation, embodied by the IgSF.

The remaining two Ig domains of CD4—D3 and D4—have also been crystallized (albeit the rat CD4 homologue) and their structures elucidated by x-ray diffraction (234,235). Remarkably, the D3D4 fragment adopts a conformation resembling that of the D1D2 portion of the molecule (see Colorplate 7), embracing prior hypotheses that CD4 arose by way of duplication of a two-domain precursor (236). D3 is a larger V-type domain homologous to D1, and D4 is a smaller module reminiscent of D2. Once again, the D3 G strand is contiguous with the A strand of D4, closely approximating the two domains and limiting the flexibility between them. However, D3 does have characteristics which distinguish it from the D1 domain. First, D3 fails to conserve the intradomain disulfide bond, resulting in a "relaxed" domain, which is packed less-tightly. Second, unlike D1, the C to C′ and F to G loops of D3 are

not shortened relative to immunoglobulin V regions. Still, these loops are unlikely to mediate dimerization of D3 domains (as they did in the Fv fragment of immunoglobulin), on account of an N-linked glycosylation site on the F strand of D3's inner face that would interfere with such interactions. Recently, a recombinant soluble form of human CD4 has also been solved crystallographically (237). While consistent with previous conclusions, this report makes the added contribution of definitively establishing the D1D2 to D3D4 junction as a hinge-like region of the protein. Thus, the rod-like two domain portions of CD4 (D1D2 and D3D4) are able to bend at a point of flexion akin to the scenario for Fab–Fc bending at the antibody hinge.

Studies have also examined the means by which CD4 binds to MHC II proteins (refer back to Fig. 23). CD4 contacts the $\alpha 2$ and $\beta 2$ domains of class II molecules using a variety of residues in the D1 and D2 domains. Most evidence points to a large surface of CD4, involving both lateral faces of D1 and the F–G loop of the D2 domain, being implicated in MHC II binding (238–240). Once more the question of the stoichiometry of interaction between coreceptor and MHC is of interest. As both sides of the D1 domain appear to contact the class II molecule, the prospect of a bivalent complex (one CD4 protein with two MHC II molecules) is once again at issue. Given that crystal studies of class II proteins have demonstrated a dimeric association between MHC II molecules, this seems a plausible mode of complex formation. Contrarily, the crystal of soluble human D1–D4 (237) has revealed a homodimeric association between D4 domains of CD4. This implies that—as had been proposed by others (241)—opposite faces of a CD4 dimer may interact with two separate class II molecules. Regardless of the specifics of their dimeric interactions, it is reasonable to conclude that multiple surfaces of CD4 are responsible for binding, and that the majority of these amino acids reside in the D1 domain.

CONCLUSION

Although these discussions only scratch the surface of structure–function relationships within the immunoglobulin superfamily of proteins, it is hoped that this chapter has served to introduce the reader to the inherent utility—and exquisite beauty—of the immunoglobulin domain as both an evolutionary tool and molecular motif. While antibody proteins have been structurally characterized and functionally probed to an unparalleled degree by the concerted and persistent efforts of the scientific community, the continued emergence of unexpected findings indicates that a great wealth of knowledge is yet to be tapped in their inquiry. Moreover, the TCR and MHC IgSF proteins occupy an even greater role in terms of immune system functioning, and their investigations have been fruitful fields for study, as well. Other molecules, like CD4 and CD8, while not the focus of attention for the prolonged duration that has been the case for immunoglobulin, have nonetheless seen seminal findings in the pursuit of their understanding, and have served to broaden our comprehension of the IgSF's variety and capacity. Still others, like the antigen-receptor signaling proteins Igα, Igβ, and CD3 subunits, or the co-stimulatory molecules B7-1 (CD80), B7-2 (CD86), CTLA-4, and CD28 have only recently been described in detail and will no doubt be centers of concentration in the immediate future. Given the unequivocal fact that new and exciting IgSF members are yet to be discovered, when one considers the pervasiveness of this class of proteins in immunology—and in biology as a whole—it is perhaps accurate to surmise that the study of the immunoglobulin superfamily is still only in its infancy.

REFERENCES

1. Dreyer WJ, Bennett JC. The molecular basis of antibody formation: a paradox. *Proc Natl Acad Sci USA* 1965;54:864–869.
2. Williams AF, Barclay AN. The immunoglobulin superfamily—domains for cell surface recognition. *Annu Rev Immunol* 1988;6:381–405.
3. Hunkapiller T, Hood L. Diversity of the immunoglobulin superfamily. *Adv Immunol* 1989;44:1–63.
4. Brummendorf T, Rathjen FG. Cell adhesion molecules 1: immunoglobulin superfamily. *Protein Profile* 1995;2:963–1108.
5. Hilschmann N, Craig LC. Amino acid sequence studies with Bence-Jones proteins. *Proc Natl Acad Sci USA* 1965;53:1403–1409.
6. Silverton EW, Navia MA, Davis DR. Three-dimensional structure of an intact human immunoglobulin. *Proc Natl Acad Sci USA* 1977;74:5140–5144.

6a. Carayannopoulous L, Capra JD. Immunoglobulins: structure and function. In: Paul WE, ed. *Fundamental Immunology,* 3rd ed. New York: Raven Press Ltd., 1993.

7. Hill RL, Delaney R, Fellows RE, Lebovitz HE. The evolutionary origins of the immunoglobulins. *Proc Natl Acad Sci USA* 1966;56:1762–1769.
8. Edmundson AB, Ely KR, Abola EE, Schiffer M, Panagiotopoulos N. Rotational allomerism and divergent evolution of domains in Ig light chains. *Biochemistry* 1975;14:3953–3961.
9. Davies DR, Padlan EA, Segal DM. Immunoglobulin structures at high resolution. *Annu Rev Biochem* 1975;44:639–667.
10. Amzel LM, Poljak RJ. Three dimensional structure of immunoglobulins. *Annu Rev Biochem* 1979;48:961–997.
11. Deisenhofer J, Colman PM, Epp O, Huber R. Crystallographic structural studies of a human Fc fragment. *Hoppe-Zeylers Z Physiol Chem* 1976;357:1421–1434.
12. Deisenhofer J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9 and 2.8 Å resolution. *Biochemistry* 1981;20:2361–2370.
13. Federation of American Society for Experimental Biology Symposium, "Two Genes, One Polypeptide Chain." *Fed Proc* 1972;31:176–209.
14. Porter RR. Separation and isolation of fractions of rabbit γ-globulin containing the antibody and antigenic combining sites. *Nature* 1958;182:670–671.
15. Nisonoff A, Wissler FC, Lipman LN. Properties of the major component of a peptic digest of rabbit antibody. *Science* 1960;132:1770–1771.
16. Palmer JL, Nisonoff A. Dissociation of rabbit γ-globulin into half-molecules after reduction of one labile disulfide bond. *Biochemistry* 1964;3:863–869.
17. Fleischman JB, Porter RR, Press E. The arrangement of the peptide chains in γ-globulin. *Biochem J* 1963;88:220–228.
18. Nisonoff A, Thorbecke GJ. Immunochemistry. *Annu Rev Biochem* 1964;33:355–402.
19. Natvig JB, Kunkel HG. Human Igs: classes, subclasses, genetic variants, and idiotypes. *Adv Immunol* 1973;16:1–59.
20. Nisonoff A, Hopper JE, Spring SB. *The antibody molecule.* New York: Academic Press, 1975.
21. Kindt TJ, Capra JD. *The antibody enigma.* New York: Plenum Press, 1984.
22. Hasemann CA, Capra JD. Immunoglobulins: structure and function. In: Paul WE, ed. *Fundamental immunology,* 2nd ed. New York: Raven Press, 1989.
23. Weir DM, ed. *Handbook of experimental immunology,* vol. 3, *Genetics & molecular immunology.* London: Blackwell Scientific Press, 1986.
24. Williams RC Jr, Kunkel HG, Capra JD. Antigenic specificities related to the cold agglutinin activity of gamma M globulins. *Science* 1968;161:379–381.

24a. Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. *Sequences of proteins of immunological interest.* Washington, DC: US Department of Health and Human Services, 1991.

25. Wu TT, Kabat EA. An analysis of the sequences of the variable regions of Bence-Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med* 1970;132:211–250.
26. Capra JD, Kehoe JM. Hypervariable regions, idiotypy, and the antibody-combining site. *Adv Immunol* 1975;20:1–40.
27. Colman PM. Structure of antibody-antigen complexes: implications for immune recognition. *Adv Immunol* 1988;43:99–132.
28. Chothia C, Novotny J, Bruccoleri R, Karplus M. Domain association in immunoglobulin molecules. *J Mol Biol* 1985;186:651–663.

28a. Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol* 1987;196:901–917.

29. Tonegawa S. Somatic generation of antibody diversity. *Nature* 1983;302:575–581.
30. Meier JT, Lewis SM. P nucleotides in V(D)J recombination: a fine-structure analysis. *Mol Cell Biol* 1993;13(2):1078–1092.
31. Meek KD, Hasemann CA, Capra JD. Novel rearrangements at the immunoglobulin D locus. *J Exp Med* 1989;170:39–57.
32. McKean D, Huppi K, Bell M, Staudt L, Gerhard W, Weigert M. Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc Natl Acad Sci USA* 1984;81:3180–3184.

33. Neuberger MS, Milstein C. Somatic hypermutation. *Curr Opin Immunol.* 1995; 7:248–254.
34. Croce CM, Shander M, Martinis J, et al. Chromosomal location of the genes for human immunoglobulin heavy chains. *Proc Natl Acad Sci USA* 1979;76: 3416–3419.
35. Malcolm S, Barton P, Murphy C, Ferguson-Smith MA, Bentley DL, Rabbitts TH. Localization of human immunoglobulin kappa light chain variable region genes to the short arm of chromosome 2 by *in situ* hybridization. *Proc Natl Acad Sci USA* 1982;79(16):4957–4961.
36. Erickson J, Martinis J, Croce CM. Assignment of the genes for human λ immunoglobulin chains to chromosome 22. *Nature* 1981;294:173–175.
37. D'Eustachio P, Pravtcheva D, Marcu K, Ruddle FH. Chromosomal location of the structural gene cluster encoding murine immunoglobulin heavy chains. *J Exp Med* 1980;151(6):1545–1550.
38. Swan D, D'Eustachio P, Leinwand L, Seidman J, Keithley D, Ruddle FH. Chromosomal assignment of the mouse κ light chain genes. *Proc Natl Acad Sci USA* 1979;76(6):2735–2739.
39. D'Eustachio P, Bothwell AL, Takaro TK, Baltimore D, Ruddle FH. Chromosomal location of structural genes encoding murine Ig λ light chains. *J Exp Med* 1981;153(4):793–800.
40. Schroeder HW Jr, Hillson JL, Perlmutter RM. Structure and evolution of mammalian V_H families: *Int Immunol* 1990;2:41–50.
41. Kirkham PM, Mortari F, Newton JA, Schroeder HW Jr. Immunoglobulin V_H clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO J* 1992;11:603–609.
42. Kirkham PM, Schroeder HW Jr. Antibody structure and the evolution of immunoglobulin V gene segments. *Semin Immunol* 1994;6:347–360.
43. Kirkham PM, Elgavish RA, Schroeder HW Jr. Structure and evolution of mammalian $V_κ$ families. *J Immunol* 1993;150:151a.
44. Kroemer G, Helmberg A, Bernot A, Auffray C, Kofler R. Evolutionary relationship between human and mouse immunoglobulin κ light chain variable region genes. *Immunogenetics* 1991;33:42–49.
45. Pascual V, Capra JD. Human immunoglobulin heavy-chain variable region genes: organization, polymorphism, and expression. *Adv Immunol* 1991;49: 1–74.
46. Zachau HG. The immunoglobulin κ locus—or—what has been learned from looking closely at one-tenth of a percent of the human genome. *Gene* 1993;135: 167–173.
47. Frippiat JP, Williams SC, Tomlinson IM, et al. Organization of the human immunoglobulin λ light-chain locus on chromosome 22q11.2. *Hum Mol Genet* 1995;4(6):983–991.
48. Cook GP, Tomlinson IM, Walter G, et al. A map of the human immunoglobulin V_H locus completed by analysis of the telomeric region of chromosome 14q. *Nat Genet* 1994;7(2):162–168.
49. Matsuda F, Shin EK, Nagaoka H, et al. Structure and physical map of 64 variable segments in the 3' 0.8 megabase region of the human immunoglobulin heavy chain locus. *Nat Genet* 1993;3:88–94.
50. Matsuda F, Honjo T. Organization of the human immunoglobulin heavy chain locus. *Adv Immunol* 1996;62:1–29.
51. Hsu E. The variation in immunoglobulin heavy chain C regions in evolution. *Semin Immunol* 1994;6:383–391.
52. Davis AC, Roux KH, Pursey J, Shulman MJ. Intermolecular disulfide bonding in IgM: effects of replacing cysteine residues in the μ heavy chain. *EMBO J* 1989;8:25–19–2526.
53. Mestecky J, Russell MW. IgA subclasses. *Monogr Allergy* 1986;19:277–301.
54. Burnett RC, Hanly WC, Zhai SK, Knight KL. The IgA heavy chain gene family in rabbit: cloning and sequence analysis of 13 Cα genes. *EMBO J* 1989;8: 4041–4047.
55. Poljak RJ. X-ray diffraction studies of immunoglobulins. *Adv Immunol* 1975;21: 1–33.
56. Huck S, Fort P, Crawford DH, Lefranc MP, Lefranc G. Sequence of a human IgG3 heavy chain C region gene: comparison with the other human Cγ genes. *Nucleic Acids Res* 1986;14:1779–1789.
57. Kocher HP, Bijlenga RK, Jaton JC. Biosynthesis and structure of membrane and secretory immunoglobulins. *Mol Cell Biochem* 1982;47:11–22.
58. Rogers J, Choi E, Souza L, et al. Gene segments encoding transmembrane carboxyl termini of Ig γ chains. *Cell* 1981:19–27.

58a. Tomlinson IM, Walter G, Jones PT, Dear PH, Sonnhammer ELL, Winter G. The imprint of somatic hypermutation on the repertoire of human germline V genes. *J Mol Biol* 1996;256:813–817.

59. Davies DR, Padlan EA, Sheriff S. Antigen-antibody complexes. *Annu Rev Biochem* 1990;59:439–473.
60. Wright A, Tao MH, Kabat EA, Morrison SL. Antibody variable region glycosylation: position effects on antigen binding and carbohydrate structure. *EMBO J* 1991;10:2717–2723.
61. Middaugh CR, Litman GW. Atypical glycosylation of an IgG monoclonal cryo-Ig. *J Biol Chem* 1987;262:3671–3673.
62. Chothia C, Lesk AM, Tramontano A, et al. Conformations of immunoglobulin hypervariable regions. *Nature* 1989;342:877–883.
63. Strong RK, Campbell R, Rose DR, Petsko GA, Sharon J, Margolies MN. Three-dimensional structure of murine anti-p-azophenylarsonate Fab 36–71. *Biochemistry* 1991;30:3739–3748.
64. Guddat LW, Herron JN, Edmundson AB. Three-dimensional structure of a human immunoglobulin with a hinge deletion. *Proc Natl Acad Sci USA* 1993;90:4271–4275.
65. Padlan EA, Davies DR. A model of the Fc of IgE. *Mol Immunol* 1986;10: 1063–1075.
66. Carayannopoulos L, Hexham JM, Capra JD. Localization of the binding site for the monocyte IgA-Fc receptor (CD89) to the domain boundary between $C_α2$ and $C_α3$ in human IgA1. *J Exp Med* 1996;183:1579–1586.

66a. Putnam FW. *The plasma proteins: structure, function, and genetic control,* vol 3, 2nd ed. New York: Academic Press, 1977.

67. Grey HM, Abel CA, Yount WJ, Kunkel HG. A subclass of human γA-globulins (γ-A2) which lacks the disulfide bonds linking heavy and light chains. *J Exp Med* 1968;128:1223–1236.
68. Saul FA, Amzel LM, Poljak RJ. Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin *New* at 2.0 A resolution. *J Biol Chem* 1978;253:585–597.
69. Padlan EA. Anatomy of the antibody molecule. *Mol Immunol* 1994;31: 169–217.
70. Davies DR, Sheriff S, Padlan E. Comparative study of two Fab-lysozyme crystal structures. *Cold Spring Harb Symp Quant Biol* 1989;54:233–238.
71. Schumaker VN, Phillips ML, Hansen DC. Dynamic aspects of antibody structure. *Mol Immunol* 1991;28:1347–1360.
72. Wade RH, Tavequ JC, Lamy JN. Concerning the axial flexibility of the Fab regions of IgG. *J Mol Biol* 1989;206:349–356.
73. Dangl JL, Wensel TG, Morrison SL, Streyer L, Herzenberg LA, Oi VT. Segmental flexibility and complement fixation of genetically engineered chimeric human, rabbit and mouse Abs. *EMBO J* 1988;7:1989–1994.
74. Spiegelberg HL. The structure and biology of human IgD. *Immunol Rev* 1977; 37:3–23.
75. Putnam FW, Liu YS, Low TL. Primary structure of a human IgA1 immunoglobulin. *J Biol Chem* 1979;254:2865–2874.
76. Plaut AG, Wistar R Jr, Capra JD. Differential susceptibility of human IgA immunoglobulins to streptococcal IgA protease. *J Clin Invest* 1974;54: 1295–1300.
77. Plaut AG, Gilbert JV, Artenstein MS, Capra JD. *Neisseriagonorrhoeae* and *Neisseria meningitidis:* extracellular enzyme cleaves human IgA. *Science* 1975;190: 1103–1105.
78. Dwek R. Glycobiology—towards understanding the function of sugars. *Biochem Soc Trans* 1995;23:1–25.
79. Jenkins M, Curling E. Glycosylation of recombinant proteins: problems and prospects. *Enzyme Microb Technol* 1995;16:354–364.
80. Schaffner G, Haase M, Geiss S. Criteria for investigation of the product equivalence of monoclonal antibodies for therapeutic and *in vivo* diagnostic use in case of change in manufacturing process. *Biologicals* 1995;23:253–259.
81. Beale D, Feinstein A. Structure and function of the C regions of immunoglobulins. *Q Rev Biophys* 1976;9:135–180.

81a. Jefferis R, Lund J. Glycosylation of antibody molecules: structural and functional significance. *Chem Immunol* 1997;65:111–128.

82. Tsuchiya N, Endo T, Matsuta K, et al. Effects of galactose depletion from oligosaccharide chains on immunological activities of human IgG. *J Rheumatol* 1989;16:285–290.
83. Tao MH, Morrison SL. Studies of aglycosylated chimeric mouse-human IgG. *J Immunol* 1989;143:2595–2601.
84. Wright A, Morrison SL. Effect of altered C_H2-associated carbohydrate structure on the functional properties and *in vivo* fate of human IgG1. *J Exp Med* 1994; 180:1087–1096.
85. Lund J, Winter G, Jones PT, et al. Human FcγRI and FcγRII interact with distinct but overlapping binding sites on the human IgG. *J Immunol* 1991;147: 2657–2662.
86. Canfield SM, Morrison SL. The binding affinity of human IgG for its high affinity Fc receptor is determined by multiple amino acids in the C_H2 domain and is modulated by the hinge region. *J Exp Med* 1991;173:1483–1491.
87. Jefferis R, Lund J. Goodall M. Recognition sites on human IgG for Fcγ receptors: the role of glycosylation. *Immunol Lett* 1995;44:111–117.

87a. Davis AC, Roux KH, Shulman MJ. On the structure of polymeric IgM. *Eur J Immunol* 1988;18:1001–1008.

88. Mole JE, Bhown AS, Bennett JC. Primary structure of human J chain: alignment of peptides from chemical and enzymatic hydrolyses. *Biochemistry* 1977;16: 3507–3513.
89. Koshland ME. The coming of age of the immunoglobulin J chain. *Annu Rev Immunol* 1985;3:425–453.
90. Max EE, Korsmeyer SJ. Human J chain gene. *J Exp Med* 1985;161:832–849.
91. Cann GM, Zaritsky A, Koshland ME. Primary structure of the immunoglobulin J chain from the mouse. *Proc Natl Acad Sci USA* 1982;79:6656–6660.
92. Hughes GJ, Frutiger S, Paquet N, Jaton JC. The amino acid sequence of rabbit J chain in secretory IgA. *Biochem J* 1990;271:641–647.
93. Mikoryak CA, Margolies MN, Steiner LA. J chain in *Rana catesbeiana* high molecular weight Ig. *J Immunol* 1988;140:4279–4285.
94. Takahashi T, Iwase T, Takenouchi N, et al. The joining (J) chain is present in invertebrates that do not express immunoglobulins. *Proc Natl Acad Sci USA* 1996;93:1886–1891.
95. Zikan J, Novotny J, Trapane TL, et al. Secondary structure of the immunoglobulin J chain. *Proc Natl Acad Sci USA* 1985;82:5905–5909.
96. Davis AC, Collins C, Yoshimura MI, D'Agostaro G, Shulman MJ. Mutations of

the mouse μ chain which prevent polymer assembly. *J Immunol* 1989;143: 1352–1357.
97. Mather EL, Alt FW, Bothwell A, Baltimore D, Koshland ME. Expression of J chain RNA in cell lines representing different stages of B lymphocyte differentiation. *Cell* 1981;23:369–378.
98. Randall TD, Parkhouse RM, Corley RB. J chain synthesis and secretion of hexameric IgM is differentially regulated by LPS and IL5. *Proc Natl Acad Sci USA* 1992;89:962–966.
99. Randall TD, King LB, Corley RB. The biological effects of IgM hexamer formation. *Eur J Immunol* 1990;20:1971–1979.
100. Underdown B. Transcytosis by the receptor for polymeric immunoglobulin. In: Metzger H, ed. *Fc receptors and the action of antibodies.* Washington, DC: American Society for Microbiology, 1990.
101. Mostov KE, Blobel G. A transmembrane precursor of secretory component. *J Biol Chem* 1982;257:11816–11821.
102. Mostov KE, Friedlander M, Blobel G. The receptor for transepithelial transport of IgA and IgM contains multiple immunoglobulin-like domains. *Nature* 1984; 308:37–43.
103. Brandtzaeg P. Human secretory IgM. *Immunology* 1975;29:559–570.
104. Weicker J, Underdown BJ. A study of the association of human secretory component with IgA and IgM proteins. *J Immunol* 1975;114:1337–1344.
105. Mach J-P. *In vitro* combination of human and bovine free secretory component with IgA of various species. *Nature* 1970;228:1278–1282.
106. Bienenstock J, Perey DY, Gauldie J, Underdown BJ. Chicken γA: physiochemical and immunochemical characteristics. *J Immunol* 1973;110:524–533.
107. Fallgreen-Gebauer E, Bebauer W, Bastian A, et al. The covalent linkage of secretory component to IgA. *Biol Chem Hoppe Seyler* 1993;374:1023–1028.
108. Bakos MA, Kurosky A, Goldblum RM. Characterization of a critical binding site for human polymeric immunoglobulin on secretory component. *J Immunol* 1991;147:3419–3426.
109. Hasemann C, Capra JD. Mutational analysis of arsonate binding by a CRIA+ antibody. *J Biol Chem* 1991;266:7626–7632.
110. Amit AG, Mariuzza RA, Phillips SEV, Poljak RJ. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science* 1986;233:747–753.
110a. Davies DR, Sheriff S, Padlan EA, Silverton EW, Cohen GH, Smith-Gill SJ. Three-dimensional structures of two Fab complexes with lysozyme. In: Smith-Gill SJ, Sercarz EE, eds. *The immune response to structurally define proteins: the lysozyme model.* Schenectady, NY: Adenine Press, 1989.
111. Padlan EA. On the nature of antibody combining sites: unusual structural features that may confer on these sites an enhanced capacity for binding ligands. *Proteins* 1990;7:112–124.
112. Mian IS, Bradwell AR, Olson AJ. Structure, function and properties of antibody binding sites. *J Mol Biol* 1991;217:133–151.
113. Mariuzza RA, Poljak RJ, Schwarz FP. The energetics of antigen-antibody binding. *Res Immunol* 1994;145:70–72.
114. Braden BC, Poljak RJ. Structural features of the reactions between antibodies and protein antigens. *FASEB J* 1995;9:9–16.
115. Herman A, Kappler JW, Marrack P, Pullen AM. Superantigens: mechanism of T-cell stimulation and role in immune responses. *Annu Rev Immunol* 1991;9:745–772.
116. Pascual V, Capra JD. B-cell superantigens? *Curr Biol* 1991;1:315.
117. Silverman GJ. Superantigens and spectrum of unconventional B-cell antigens. *The Immunologist* 1994;2:51–57.
118. Biguzzi S. Fcγ-like determinants on immunoglobulin variable regions: identification by staphylococcal protein A. *Scand J Immunol* 1982;15:605–618.
119. Sasso EH, Silverman GJ, Mannik M. Human IgM molecules that bind staphylococcal protein A contain V_HIII H chains. *J Immunol* 1989;142:2778–2783.
120. Potter KN, Li Y, Capra JD. Staphylococcal protein A simultaneously interacts with FR1, CDR2, and FR3 on human V_H3-encoded Igs. *J Immunol* 1996;157: 2982–2988.
121. Berberian L, Goodglick L, Kipps TJ, Braun J. Immunoglobulin V_H3 gene products: natural ligands for HIV gp120. *Science* 1993;261:1588–1591.
122. Lenert P, Kroon D, Spiegleberg H, Golub ES, Zanetti M. Human CD4 binds immunoglobulins. *Science* 1990;248:1639–1643.
123. Corper AL, Sohi MK, Bonagura VR, et al. Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction. *Nat Struct Biol* 1997;4:374–381.
123a. Simard J, Mak TW. Immunoglobulins and the B cell receptor. In: *Genetic basis of immune responses* (in press).
124. Hombach J, Lottspeich F, Reth M. Identification of the genes encoding the IgM-α and Ig-β components of the IgM antigen receptor complex by amino-terminal sequencing. *Eur J Immunol* 1990;20:2795–2799.
125. Williams GT, Venkitaraman AR, Gilmore DJ, Neuberger MS. The sequence of the μ transmembrane segment determines the tissue specificity of the transport of IgM to the cell surface. *J Exp Med* 1990;171:947–952.
126. Karasuyama H, Kudo A, Melchers F. The proteins encoded by the VpreB and λ5 pre-B cell-specific genes can associate with each other and with μ heavy chain. *J Exp Med* 1990;172:969–972.
127. Tsubata T, Reth M. The products of pre-B cell-specific genes (λ5 and VpreB) and the immunoglobulin μ chain form a complex that is transported onto the cell surface. *J Exp Med* 1990;172:973–976.
128. Arya S, Chen F, Spycher S, Isenman DE, Shulman MJ, Painter RH. Mapping of amino acid residues in the C_H3 domain of mouse IgM important in macromolecular assembly and complement-dependent cytolysis. *J Immunol* 1994;152: 1206–1212.
129. Wright JF, Shulman MJ, Isenman DE, Painter RH. C1 binding by mouse IgM. *J Biol Chem* 1990;265:10506–10513.
130. Du Pasquier L, Wabl MR. Antibody diversity in amphibians: inheritance of isoelectric focusing antibody patterns in isogenic frogs. *Eur J Immunol* 1978;8: 428–433.
131. Lin LC, Putnam FW. Primary structure of the Fc region of human IgD: implications of evolutionary origin and biologic function. *Proc Natl Acad Sci USA* 1981;78:504–508.
132. Nitschke L, Kosco MH, Kohler G, Lamers MC. IgD-deficient mice can mount normal immune responses to thymus-independent and -dependent antigens. *Proc Natl Acad Sci USA* 1993;90:1887–1891.
133. Roes J, Rajewsky K. Immunoglobulin D (IgD)-deficient mice reveal an auxiliary receptor function for IgD in antigen-mediated recruitment of B cells. *J Exp Med* 1993;177:45–55.
134. Forster I, Vieira P, Rajewsky K. Flow cytometric analysis of cell proliferation dynamics in the B cell compartment of the mouse. *Int Immunol* 1989;1(4): 321–331.
135. Black SJ, van der Loo W, Loken MR, Herzenberg LA. Expression of IgD by murine lymphocytes. Loss of surface IgD indicates maturation of memory B cells. *J Exp Med* 1978;147:984–996.
136. McHeyzer-Williams MG, Nossal GJ, Lalor PA. Molecular characterization of single memory B cells. *Nature* 1991;350:502–505.
137. Havran WL, DiGiusto DL, Cambier JC. mIgM:mIgD ratios on B cells: mean mIgD expression exceeds mIgM by 10-fold on most splenic B cells. *J Immunol* 1984;132:1712–1716.
138. Venkitaraman AR, Williams GT, Dariavach P, Neuberger MS. The B-cell antigen receptor of the five immunoglobulin classes. *Nature* 1991;352:777–781.
139. Nezlin R. Internal movements in immunoglobulin molecules. *Adv Immunol* 1990;48:1–40.
140. Simard J, Mak TW. Immunoglobulins and the B cell receptor. In: *Genetic basis of immune responses* (in press).
141. Brink R, Goodnow CC, Crosbie J, et al. IgM and IgD antigen receptors are both capable of mediating B lymphocyte activation, deletion, or anergy after interaction with specific antigen. *J Exp Med* 1992;176:991–1005.
142. Morris SC, Lees A, Finkelman FD. *In vivo* activation of naïve T cells by antigen-presenting B cells. *J Immunol* 1994;152:3777–3785.
143. Brink R, Goodnow CC, Basten A. IgD expression on B cells is more efficient than IgM but both receptors are functionally equivalent in up-regulating CD80/CD86 co-stimulatory molecules. *Eur J Immunol* 1995;25:1980–1984.
144. Swenson CD, Van Vollenhoven RF, Xue B, Siskind GW, Thorbecke GJ, Coico RF. Physiology of IgD. *Eur J Immunol* 1988;18:13–20.
145. Finkelman FD, Snapper CM, Mountz JD, Katona IM. Polyclonal activation of the murine immune system by a goat antibody to mouse IgD. *J Immunol* 1987;138:2826–2830.
146. Liu Y-J, de Bouteiller O, Arpin C, et al. Normal human IgD+/IgM+ germinal center B cells can express up to 80 mutations in the variable region of their IgD transcripts. *Immunity* 1996;4:603–613.
147. Billian G, Bella C, Mondiere P, Defrance T. Identification of a tonsil IgD+ B cell subset with phenotypical and functional characteristics of germinal center B cells. *Eur J Immunol* 1996;26:1712–1719.
148. Perlemutter RG, Hansburg D, Briles DE, Nicolotti RA, Davis JM. Subclass restriction of murine anti-carbohydrate antibodies. *J Immunol* 1978;121: 566–571.
149. Coutelier J-P, van der Logt JTM, Heessen FWA, Warnier G, Van Snick J. IgG2a restriction of murine antibodies elicited by viral infections. *J Exp Med* 1987;165: 64–69.
150. Papadea C, Check IJ. Human IgG and IgG subclasses: biochemical, genetic, and clinical aspects. *Crit Rev Clin Lab Sci* 1989;27:27–58.
151. Brekke OH, Michaelsen TE, Sandlie I. The structural requirements for complement activation by IgG: does it hinge on the hinge? *Immunol Today* 1995;16: 85–90.
152. Bruggemann M, Williams GT, Bindon CI, et al. Comparison of the effector functions of human immunoglobulins using a matched set of chimeric antibodies. *J Exp Med* 1987;166:1351–1361.
153. Duncan AR, Winter G. The binding site for C1q on IgG. *Nature.* 1988;332: 738–740.
154. Tao M-H, Canfield SM, Morrison SL. The differential ability of human IgG1 and IgG4 to activate complement is determined by the COOH-terminal sequence of the C_H2 domain. *J Exp Med* 1991;173:1025–1028.
155. Tan LK, Shopes RJ, Oi VT, Morrison SL. Influence of the hinge region on complement activation, C1q binding, and segmental flexibility in chimeric human immunoglobulins. *Proc Natl Acad Sci USA* 1990;87:162–166.
156. Sandlie I, Michaelsen TE. Engineering monoclonal antibodies to determine the structural requirements for complement activation and complement mediated lysis. *Mol Immunol* 1991;28:1361–1368.
157. Brekke OH, Michaelsen TE, Sandin R, Sandlie I. Activation of complement by an IgG molecule without a genetic hinge. *Nature* 1993;363:628–630.
158. Norderhaug L, Brekke OH, Bremnes B, et al. Chimeric mouse-human IgG3 antibodies with an IgG4-like hinge region induce complement-mediated lysis more efficiently than IgG3 with normal hinge. *Eur J Immunol* 1991;21:2379–2384.

159. Kimberly RP, Salmon JE, Edberg JC. Receptors for IgG. *Arthritis Rheum* 1995;38:306–314.
160. Indik ZK, Park JG, Hunter S, Schreiber AD. Structure/function relationships of Fcγ receptors in phagocytosis. *Semin Immunol* 1995;7:45–54.
161. Unkeless JC. Function and heterogeneity of human Fc receptors for IgG. *J Clin Invest* 1989;83:355–361.
162. Michaelsen TE, Aase A, Norderhaug L, Sandlie I. Antibody dependent cell-mediated cytotoxicity induced by chimeric mouse-human IgG subclasses and IgG3 antibodies with altered hinge region. *Mol Immunol* 1992;29:319–326.
163. Galon J, Bouchard C, Fridman WH, Sautes C. Ligands and biological activities of soluble Fcγ receptors. *Immunol Lett* 1995;44:175–181.
164. Saji F, Koyama M, Matsuzaki N. Current topic: human placental Fc receptors. *Placenta* 1994;15:453–466.
165. Sarmay G, Lund J, Rozsnyay Z, Gergely J, Jefferis R. Mapping and comparison of the interaction sites on the Fc region of IgG responsible for triggering ADCC through different types of human Fcγ receptor. *Mol Immunol* 1992;29: 633–639.
166. Canfield SM, Morrison SL. The binding affinity of human IgG for its high affinity Fc receptor is determined by multiple amino acids in the C_H2 domain and is modulated by the hinge region. *J Exp Med* 1991;173:1483–1491.
167. Tomasi TB, Grey HM. Structure and function of IgA. *Prog Allergy* 1972;16: 81–213.
168. Lehner T, Bergmeier LA, Tao L, et al. Targeted lymph node immunization with simian immunodeficiency virus p27 antigen to elicit genital, rectal, and urinary immune responses in nonhuman primates. *J Immunol* 1994;153:1858–1868.
169. Mestecky J, McGhee JR. Immunoglobulin A (IgA): molecular and cellular interactions involved in IgA biosynthesis and immune response. *Adv Immunol* 1987;40:153–245.
170. Qiu JZ, Brackee GP, Plaut AG, Sneller MC, Strober W. Analysis of the specificity of bacterial IgA proteases by a comparative study of ape serum IgAs as substrates. *Infect Immun* 1996;64:933–937.
171. Wang B, Rieger A, Kilgus O, et al. Human epidermal Langerhans cells from normal human skin bind monomeric IgE via FcεRI. *J Exp Med* 1992;175: 1353–1365.
172. Bieber T, de la Salle H, Wollenberg A, et al. Human epidermal Langerhans cells express the high affinity receptor for IgE (FcεRI). *J Exp Med* 1992;175: 1285–1290.
173. Gounni AS, Lamkhioued B, Ochiai K, et al. High-affinity IgE receptor on eosinophils is involved in defense against parasites. *Nature* 1994;367:183–186.
174. Hashimoto S, Koh K, Tomita Y, et al. TNF-α regulates IL-4-induced FcεRII/CD23 gene expression and soluble FcεRII release by human monocytes. *Int Immunol* 1995;7:705–713.
175. Helm B, Marsh P, Vercelli D, Padlan E, Gould H, Geha R. The mast cell binding site on human immunoglobulin E. *Nature* 1988;331:180–183.
176. Nissim. A, Eshhar Z. The human mast cell receptor binding site maps to the third constant domain of IgE. *Mol Immunol* 1992;29:1065–1072.
177. Edelman GM, Cunningham BA, Gall WE, Gottlieb PD, Rutishauser U, Waxdal MJ. The covalent structure of an entire γG immunoglobulin molecule. *Proc Natl Acad Sci USA* 1969;63:78–85.
178. Brummendorf T, Rathjen FG. Structure/function relationships of axon-associated adhesion receptors of the immunoglobulin superfamily. *Curr Opin Neurobiol* 1996;6:584–593.
179. Edelman GM. CAMs and Igs: cell adhesion and the evolutionary origins of immunity. *Immunol Rev* 1987;100:11–45.
180. Hultgren SJ, Jacob-Dubuisson F, Jones CH, Branden CI. PapD and superfamily of periplasmic immunoglobulin-like pilus chaperones. *Adv Protein Chem* 1993; 44:99–123.
181. Bateman A, Eddy SR, Chothia C. Members of the immunoglobulin superfamily in bacteria. *Protein Sci* 1996;5:1939–1941.
182. Wojciechowicz D, Lu CF, Kurjan J, Lipke PN. Cell surface anchorage and ligand-binding domains of the *Saccharomyces cerevisiae* cell adhesion protein alpha-agglutinin, a member of the IgSF. *Mol Cell Biol* 1993;13:2554–2563.
183. Ohno S. The origin of Igs and T-cell receptors is likely to be the cell death sensor of macrophages. *Res Immunol* 1996;147:247–252.
184. Noegel A, Gerisch G, Stadler J, Westfal M. Complete sequence and transcript regulation of a cell adhesion protein from aggregating *Dictyostelium* cells. *Embo J* 1986;5:1473–1476.
185. Mokady O. Occam's Razor, invertebrate allorecognition and immunoglobulin superfamily evolution. *Res Immunol* 1996;147:241–246.
186. Lesk AM, Chothia C. Evolution of proteins formed by β-sheets. *J Mol Biol* 1982; 160:325–342.
187. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *EMBO J* 1986;5:823–826.
188. Harpaz Y, Chothia C. Many of the IgSF domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J Mol Biol* 1994;238:528–539.
189. Kolsch E, Oberbarnscheidt J, Bruner K, Heuer J. The Fc receptor: its role in the transmission of differentiation signals. *Immunol Rev* 1980;49:61–78.
190. Fridman WM. Fc receptors and immunoglobulin binding factors. *FASEB J* 1991;5:2684–2690.
191. Unkeless JC, Scigliano, E, Freedman VH. Structure and function of human and murine receptors for IgG. *Annu Rev Immunol* 1988;6:251–281.
192. Sandor M, Lynch R. The biology and pathology of Fc receptors. *J Clin Immunol* 1993;13:237–246.
193. Keegan AD, Paul WE. Multichain immune recognition receptors: similarities in structure and signaling pathways. *Immunol Today* 1992;13:63–68.
194. Lanier L, Yu G, Phillips J. Analysis of FcγRIII (CD16) membrane expression and association with CD3ζ and FcεRI-γ by site directed mutation. *J Immunol* 1991; 146:1571–1576.
195. Cassel DL, Keller MA, Surrey S, et al. Differential expression of FcγRIIA, FcγRIIB and FcγRIIC in hematopoietic cells: analysis of transcripts. *Mol Immunol* 1993;30:451–460.
196. Heyman B. The immune complex: possible ways of regulating the antibody response. *Immunol Today* 1990;11:310–313.
197. Vivier E, Daeron M. Immunoreceptor tyrosine-based inhibition motifs. *Immunol Today* 1997;18:286–291.
198. Hogarth PM, Hulett MD, Ierino FL, Tate B, Powell MS, Brinkworth RL. Identification of the immunoglobulin binding regions of FcγRII and FcεRI. *Immunol Rev* 1992;125:21–35.
199. Hulett MD, Witort E, Brinkworth RI, McKenzie IF, Hogarth PM. Multiple regions of human FcγRII (CD32) contribute to the binding of IgG. *J Biol Chem* 1995;270:21188–21194.
200. Letourneur O, Kennedy IC, Brini AT, Ortaldo JR, O'Shea JJ, Kinet JP. Characterization of the family of dimers associated with Fc receptors (FcεRI and FcγRIII). *J Immunol* 1991;147:2652–2656.
201. Selvaraj P, Rosse WF, Silber R, Springer TA. The major Fc receptor in blood has a phosphatidylinositol anchor and is deficient in paroxysmal nocturnal haemoglobinuria. *Nature* 1988;333:565–567.
202. Metzger H. The receptor with high affinity for IgE. *Immunol Rev* 1992;125:38–48.
203. Maurer D, Fiebiger E, Reininger B, et al. Expression of functional high affinity IgE receptors (FcεRI) on monocytes of atopic individuals. *J Exp Med* 1994;179: 745–750.
204. Dombrowicz D, Flamand V, Brigman K, et al. Abolition of anaphylaxis by targeted disruption of the high affinity IgE receptor alpha chain gene. *Cell* 1993;75:969–976.
205. Takai T, Li M, Sylvestre D, et al. FcR γ chain deletion results in pleiotropic effector cell defects. *Cell* 1994;76:519–529.
206. Clynes R, Ravetch J. Cytotoxic antibodies trigger inflammation through Fc receptors. *Immunity* 1995;3:21–26.
207. Leahy DJ. A structural view of CD4 and CD8. *FASEB J* 1995;9:17–25.
208. Rosenstein Y, Ratnofshy S, Burakoff SJ, Herrmann SH. Direct evidence for binding of CD8 to HLA class I antigens. *J Exp Med* 1989;169:149–160.
209. Weber S, Karjalainen K. Mouse CD4 binds MHC class II with extremely low affinity. *Int Immunol* 1993;5:695–698.
210. Littman DR. The structure of the CD4 and CD8 genes. *Annu Rev Immunol* 1987; 5:561–584.
211. Veillette A, Bookman MA, Horak EM, Bolen JB. The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase $p56^{lck}$. *Cell* 1988;55:301–308.
212. Wallace VA, Penninger J, Mak TW. CD4, CD8 and tyrosine kinases in thymic selection. *Curr Opin Immunol* 1983;5:235–240.
213. Janeway CA. The T cell receptor as a multicomponent signaling machine: CD4/CD8 coreceptors, and CD45 in T cell activation. *Annu Rev Immunol* 1992;10: 645–674.
214. Littman DR, Thomas Y, Maddon PJ, Chess L, Axel R. The isolation and sequence of the gene encoding T8: a molecule defining functional classes of T lymphocytes. *Cell* 1985;40:237–246.
215. Norment AM, Littman DR. A second subunit of CD8 is expressed in human T cells. *EMBO J* 1988;7:3433–3439.
216. Shiue L, Gorman SD, Parnes JR. A second chain of human CD8 is expressed on peripheral blood lymphocytes. *J Exp Med* 1988;168:1993–2005.
217. Wheeler CJ, von Hoegen P, Parnes JR. An immunological role for the CD8β-chain. *Nature* 1992;357:247–249.
218. Nakayama K, Nakayama K, Negishi I, et al. Requirement for CD8β chain in positive selection of CD8-lineage T cells. *Science* 1994;263:1131–1133.
219. Renard V, Romero P, Vivier E, Malissen B, Luescher IF. CD8β increases CD8 coreceptor function and participation in TCR-ligand binding. *J Exp Med* 1996; 184:2439–2444.
220. Irie HY, Ravichandran KS, Burakoff SJ. CD8β chain influences CD8α chain-associated Lck kinase activity. *J Exp Med* 1995;181:1267–1273.
221. Nakayama K, Kawachi Y, Tokito S, et al. Recent duplication of the two human CD8 β-chain genes. *J Immunol* 1992;148:1919–1927.
222. DiSanto JP, Smith D, de Bruin D, Lacy E, Flomenberg N. Transcriptional diversity at the duplicated human CD8β loci. *Eur J Immunol* 1993;23:320–326.
223. Leahy DJ, Axel R, Hendrickson WA. Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. *Cell* 1992;68:1145–1162.
224. Kirszbaum L, Sharpe JA, Goss N, Lahnstein J, Walker ID. The α-chain of murine CD8 lacks an invariant Ig-like disulfide bond but contains a unique intrachain loop instead. *J Immunol* 1989;142:3931–3936.
225. Classon BJ, Brown MH, Garnett D, et al. The hinge region of the CD8α chain: structure, antigenicity, and utility in expression of IgSF domains. *Int Immunol* 1992;4:215–225.
226. Pascale MC, Erra MC, Malagolini N, Serafini-Cessi F, Leone A, Bonatti S. Post-translational processing of an O-glycosylated protein, the human CD8 glycopro-

tein, during the intracellular transport to the plasma membrane. *J Biol Chem* 1992;267:25196–25201.
227. Boursier JP, Alcover A, Herve F, Laisney I, Acuto O. Evidence for an extended structure of the T-cell co-receptor CD8α as deduced from the hydrodynamic properties of soluble forms of the extracellular region. *J Biol Chem* 1993;268:2013–2020.
228. Sanders SK, Fox RO, Kavathas P. Mutations in CD8 that affect interactions with HLA class I and monoclonal anti-CD8 antibodies. *J Exp Med* 1991;174:371–379.
229. Giblin PA, Leahy DJ, Mennone J, Kavathas PB. The role of charge and multiple faces of the CD8 α/α homodimer in binding to MHC class I molecules: support for a bivalent model. *Proc Natl Acad Sci USA* 1994;91:1716–1720.
230. Gao GF, Tormo J, Gerth UC, et al. Crystal structure of the complex between human CD8α/α and HLA-A2. *Nature* 1997;387:630–634.
231. Maddon PJ, Littman DR, Godfrey M, Maddon DE, Chess L, Axel R. The isolation and nucleotide sequence of a cDNA encoding the T-cell surface protein T4: a new member of the IgSF. *Cell* 1985;42:93–104.
232. Wang JH, Yau YW, Garrett TP, et al. Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature* 1990;348:411–418.
233. Ryu SE, Kwong PD, Truneh A, et al. Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature* 1990;348:419–426.
234. Brady RL, Dodson EJ, Dodson GG, et al. Crystal structure of domains 3 and 4 of rat CD4: relation to the NH2-terminal domains. *Science* 1993;260:979–983.
235. Lange G, Lewis SJ, Murshudov GN, et al. Crystal structure of an extracellular fragment of the rat CD4 receptor containing domains 3 and 4. *Structure* 1994;2:469–481.
236. Williams AF, Davis SJ, He Q, Barclay AN. Structural diversity in domains of the immunoglobulin superfamily. *Cold Spring Harb Symp Quant Biol* 1989;54:637–647.
237. Wu H, Kwong PD, Hendrickson WA. Dimeric association and segmental variability in the structure of human CD4. *Nature* 1997;387:527–530.
238. Fleury S, Lamarre D, Meloche S, et al. Mutational analysis of the interaction between CD4 and class II MHC: class II antigens contact CD4 on a surface opposite the gp120-binding site. *Cell* 1991;66:1037–1049.
239. Moebius U, Pallai P, Harrison SC, Reinherz EL. Delineation of an extended surface contact area on human CD4 involved in class II major histocompatibility complex binding. *Proc Natl Acad Sci USA* 1993;90:8259–8263.
240. Huang B, Yachou A, Fleury S, Hendrickson WA, Sekaly RP. Analysis of the contact sites on the CD4 molecule with class II MHC molecule: co-ligand versus co-receptor function. *J Immunol* 1997;158:216–225.
241. Langedijk JP, Puijk WC, van Hoorn WP, Meloen RH. Location of CD4 dimerization site explains critical role of CDR3-like region in HIV-1 infection and T-cell activation and implies a model for complex of coreceptor-MHC. *J Biol Chem* 1993;268:16875–16878.